FUNDAMENTALS OF PHYSICS

B.N. Ivanov

Mir Publishers Moscow

FUNDAMENTALS OF PHYSICS

This book deals with the basic concepts of physics at the high-school level. In addition to the traditional aspects considered in such a course, the book also contains section on modern physics (astrophysics, plasma physics, physical hydrodynamics, etc.). The distinguishing feature of the book is that the laws of macrophysics are derived from an analysis of the phenomena associated with particles and fields.

Although this is not a textbook on physics in the strict sense (it is rather intended to be a helpbook), it cannot be treated as a light reading material either. It requires a great deal of thinking, analysis, and even computations at times, on the part of the reader. The mathematical part has been kept on a low profile and care has been taken to describe the physical principles underlying phenomena. Wherever mathematical description is found to be unavoidable, a background has been provided for those not initiated into the intricate aspects of mathematical analysis.

Intended for a wide range of readers, and also as a teach-yourself book.

ISBN 5-03-000527-7

$$m\frac{\Delta v}{\Delta t}=F$$

$$x=x'+Vt$$

$$p=\sum_i p_i=const$$

$$E=T+U$$

$$F=G\frac{m_1 m_2}{r^2}$$

$$F=mg$$

$$E=\frac{m}{\sqrt{1}}$$

$$p=\frac{mv}{\sqrt{1-\frac{v}{c}}}$$

$$E^2=c^2p^2+$$

$$E=cp$$

$$\begin{cases} x=\frac{x'+}{\sqrt{1-}} \\ t=\frac{t'+}{\sqrt{1-}} \end{cases}$$

$$\Delta p_x \Delta x \propto \hbar$$

$$\Psi=\psi e^{i\varphi}$$

$$r_a=\frac{\hbar^2}{me^2}$$

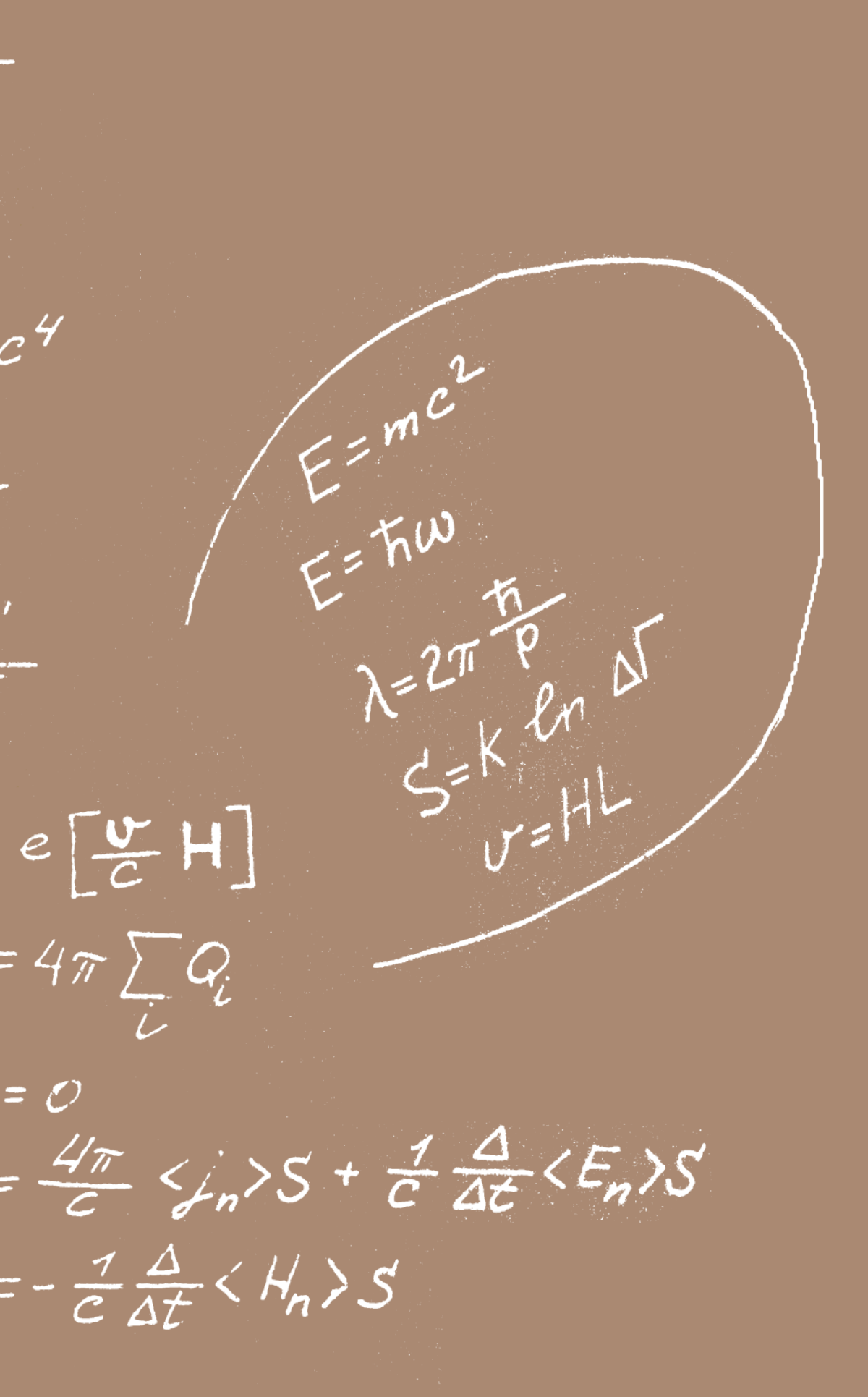
$$E_a=-\frac{me^4}{2\hbar^2}$$

$$\omega_{iK}=\frac{E_i-E_K}{\hbar}$$

$$\alpha=\frac{e^2}{\hbar c}$$

$$F=$$

$$\begin{cases} <E_n \\ <H_n \\ <H_i \\ <E_i \end{cases}$$


c^4
$E = mc^2$
$E = \hbar\omega$
$\lambda = 2\pi \frac{\hbar}{p}$
$S = k \ln \Delta\Gamma$
$V = HL$
$e\left[\frac{v}{c} H\right]$
$= 4\pi \sum_i Q_i$
$= 0$
$= \frac{4\pi}{c} \langle j_n \rangle S + \frac{1}{c} \frac{\Delta}{\Delta t} \langle E_n \rangle S$
$= -\frac{1}{c} \frac{\Delta}{\Delta t} \langle H_n \rangle S$

$$|E| = \frac{Q}{r^2}$$

$$|F| = \frac{Q_1 Q_2}{r^2}$$

$$U = \frac{Q_1 Q_2}{r}$$

$$\varphi = \frac{Q}{r}$$

$$|E| = -\frac{\Delta\varphi}{\Delta r}$$

$$P_m = \frac{Q}{2mc} L$$

$$E_1 = \frac{Q}{r^2} \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$$

$$J \approx \frac{Q^2}{c^3} a^2$$

$$T_d \sim \frac{\hbar^2}{mk} \left(\frac{N}{V}\right)^{2/3}$$

$$C_p - C_V = R$$

$$\gamma = \sigma T^4$$

$$p_F \sim \hbar \left(\frac{Ne}{V}\right)^{1/3}$$

$$F = C_p u^2 l^2$$

$$D \sim \frac{\lambda}{\rho C_p} \sim \frac{\eta}{\rho}$$

$$\frac{\sqrt{\langle (\Delta f^2) \rangle}}{\langle f \rangle} \sim \frac{1}{\sqrt{N}}$$

$$\frac{1}{T} = \frac{dS}{dE}$$

$$\omega = A e^{-E_n/(kT)}$$

$$T = \frac{E_1 - E_0}{k \ln\left(\frac{n_0}{n_1}\right)}$$

$$dE = dQ + dA$$

$$pV = RT$$

$$\frac{1}{2} m \langle v_\alpha^2 \rangle = \frac{3}{2} \kappa$$

$$\frac{dp}{dT} = \frac{q}{T(V_g -$$

$$\frac{v^2}{2} + \frac{p}{\rho} = \text{const}$$

$$Re = \frac{\rho v L}{\eta}$$

$$\begin{cases} \rho_1 v_1 = \rho_2 v_2 \\ h_1 + \frac{v_1^2}{2} = \end{cases}$$

$$\frac{\lambda}{\sigma T} \sim \left(\frac{k^2}{e^2}\right)$$

$$\sigma \sim \frac{e^2 n \tau}{m}$$

$$P_* \sim \frac{G M_*^2}{R_*^4}$$

$$R = \rho \frac{L}{S}$$

$$T_0 \lambda_{max} = const$$

$$M_* \approx \left(\frac{c\hbar}{G}\right)^{3/2} \frac{1}{m_N^2}$$

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

$$m_i = m_g$$

$$n = \sqrt{\varepsilon \mu}$$

$$J_* \sim \frac{1}{c^3} \rho_m^2 \omega^4$$

$$\varepsilon = 1 + \alpha N$$

$$\frac{n_\gamma}{n_p} \sim 10^9$$

$$\Phi_0 = \pi \frac{\hbar c}{e}$$

$$\frac{1}{H} \simeq 2 \times 10^{10} \text{ years}$$

$$\upsilon_p = \sqrt{\frac{4\pi N_e e^2}{m_e}}$$

$$D = kTb$$

$$^8B \nearrow {}^8Be^* + e^+ + \nu_e$$

$$\left.\begin{array}{l} \rho_1^2 = \rho_2 + \rho_2 \upsilon_2^2 \\ \frac{\rho_2}{\rho_1} \end{array}\right\}$$

FUNDAMENTALS OF PHYSICS

Б. Н. ИВАНОВ

ЗАКОНЫ ФИЗИКИ

«Высшая школа» Москва

B.N. Ivanov

FUNDAMENTALS OF PHYSICS

Mir
Publishers
Moscow

Translated from Russian
by R.S. Wadhwa, Ph.D.

First published 1989
Revised from the 1986 Russian edition

На английском языке

Printed in the Union of Soviet Socialist Republics

ISBN 5-03-000527-7

© Издательство «Высшая школа», 1986
© English translation, Mir Publishers, 1989

PREFACE

Over the last few years, several books have appeared for preparatory courses in colleges and universities. A student can select a book according to his requirements and taste with or without his teachers' guidance. The book being offered by the author differs from other existing books on the subject in its nontraditional approach to the course of physics. The principle underlying the preparation of this course can be summarized as follows: "From atom to matter".

What prompted the author to adopt this approach? Indeed, the creation of new materials with unusual mechanical, thermal, electrical, magnetic, and optical properties requires a microscopic approach to the problem and a clear understanding of the practical significance of the approach "from atom to matter". This means that the scientists and industrial workers engaged in fields like physical materials science, nuclear and semiconductor engineering, laser technology, and space science must be prepared to use this approach.

Modern radio and electronics engineering is becoming quantum-mechanical in its approach. Today, electrical engineers use superconducting materials. Many important biological problems including the mechanism of muscle operation, transfer of nerve signals, and the mechanism of memory can be solved only by molecular-level research. These requirements are not fully met by the courses of physics using the traditional scheme: mechanics, heat, electricity and magnetism, optics, atomic physics. Lastly, it should be remarked that the formation and teaching of physics courses on the approach "from atom to matter" is a requirement not only of the modern engineering physics practice but also of the logic of development of physics.

What will be the approach of the "natural science ideology" towards the technical progress in the next century? How should physics be taught so that a young scientist can find his way in the myriads of scientific and technical information, and independently sift this information to grasp the

main trend? An optimal solution of these questions still remains to be found.

The list of contents provides a detailed information about the material contained in this book. It can be seen that this book covers a wide range of topics.

The International System of units (SI) has been used practically throughout the book. However, since the CGS system is still widely used in research and scientific publications, students of physics must be familiar with both systems and freely move from one system to the other. The material in Chaps. 3, 4, and 7 of this book uses the CGS system.

This book is intended for those who wish to acquire a deeper knowledge of physical phenomena. It can be used by students of physics and mathematical schools, as well as by those who have finished school and are engaged in self-education. A good deal of the material may be useful to teachers delivering lectures on various topics of physics.

In conclusion, the author would like to place on record his profound gratitude to Profs. L. V. Tarasov and A. D. Sukhanov for their benevolent comments which helped a lot in the improvement of this book. Special thanks are due to T. D. Kirilina for her help in preparing the manuscript.

Author

TO THE READER

This is not a textbook, but rather a helpbook that should be used in conjunction with the standard textbooks. Nor is the book intended for a light reading; you have to use a pen and paper, think, analyze, and even compute whenever it is necessary. We shall describe physics here in the way researchers understand it today.

Physics essentially deals with the fundamental laws of nature. The progress being made at present in all branches of natural science is due, as a rule, to the introduction of physical concepts and techniques in them. This is besides the fact that a knowledge of physical sciences is essential for new industrial ventures lying at the root of technical progress. Physics is fast becoming an important element in the modern civilization.

While presenting the material in this book, I shall presume that the reader possesses a certain knowledge about things and events. Of course, he may not be always aware of the nature of these things or the causes behind the events. For this purpose, we embark upon an arduous but, hopefully, fascinating journey to the "Land of Physics". I also presume that the reader is in a position to use the mathematical apparatus. This is the language in which we shall have to converse with Nature. I shall try to provide necessary explanations wherever an advanced level of mathematical knowledge is required.

B. Ivanov

CONTENTS

Preface 5
To the Reader 7

CHAPTER 1. UNITY OF NATURE 16

1.1. Hierarchy of Natural Objects 16
1.1.1. Elementary Particles 16
1.1.2. Nuclei 20
1.1.3. Atoms and Molecules 21
1.1.4. Macroscopic Bodies 22
1.1.5. Planets 23
1.1.6. Stars. Galaxies. Universe 25
1.2. Four Types of Fundamental Interactions 26
1.2.1. Bound Systems of Objects. Interactions 26
1.2.2. Gravitational Interactions 26
1.2.3. Electromagnetic Interactions 27
1.2.4. Strong (Nuclear) Interactions 27
1.2.5. Weak Interactions 27
1.2.6. Comparative Estimates for the Intensity of All Types of Interactions 29
1.2.7. Fields and Matter 29
1.3. Space and Time 30
1.3.1. Scales of Space and Time in Nature 30
1.3.2. Homogeneity of Space and Time 31
1.3.3. Free Bodies and Inertial Motion 31
1.3.4. Inertial Reference Frames. The Relativity Principle 32

CHAPTER 2. MECHANICS OF A MATERIAL PARTICLE 34

2.1. Coordinates, Velocity, Acceleration 34
2.2. Galilean Transformations 35
2.2.1. Absolute Nature of Dimensions and Time Intervals 36
2.2.2. Relative Nature of Velocities and the Law of Their Transformation 37
2.2.3. Absolute Nature of Accelerations 37
2.3. Law of Motion in Mechanics 37
2.4. Motion of a Material Particle in a Gravitational Field 39

2.5. Momentum. Law of Momentum Conservation 42
2.6. Law of Energy Conservation. Applications and Universal Nature of Conservation Laws 43
2.6.1. Law of Energy Conservation 43
2.6.2. Applications of Conservation Laws 46
2.6.3. Universal Nature of Conservation Laws. Angular Momentum 52
2.7. Ultimate Velocity. Mechanics of High-Energy Particles 54
2.7.1. Experiments on Accelerators and Ultimate Velocity 54
2.7.2. Lorentz Transformations 55
2.7.3. Relativistic Energy and Momentum 58
2.7.4. Role of Relativistic Constant c in Physics 61

CHAPTER 3. ELECTROMAGNETIC FIELD 63
3.1. Electric Charge 63
3.2. Method of Field Investigation 64
3.2.1. Equation of Motion of a Charge in a Field 64
3.2.2. Laws of Field Transformation 64
3.3. Laws of Electromagnetic Field 66
3.3.1. New Objects and New Mathematics 66
3.3.2. First Field Equation. Relation Between Electric Field and Electric Charge 67
3.3.3. Second Field Equation. Absence of Magnetic Charges 68
3.3.4. Third Field Equation. Relation Between Current and "Something" with a Vortex Magnetic Field 68
3.3.5. Fourth Field Equation. Relation Between a Varying Magnetic Field and a Vortex Electric Field 71
3.3.6. Additional Analysis of the Third Field Equation. Relation Between a Varying Electric Field and a Vortex Magnetic Field 72
3.3.7. Maxwell's Field Equations 73
3.4. Constant Electric Field 74
3.4.1. Field of a Stationary Point Charge 74
3.4.2. Field of Charges Distributed over a Sphere, Line or Plane Surface 74
3.4.3. Electrostatic Energy of Charges. Field Potential 77
3.4.4. Field of a Dipole. Charge-Dipole and Dipole-Dipole Interactions 80
3.5. Constant Magnetic Field 82
3.5.1. Magnetic Field of a Direct Current 82
3.5.2. Magnetic Field of a Current Surface 82
3.5.3. Magnetic Moment and Its Relation with Mechanical (Angular) Momentum 83

3.6. Motion of Charges in a Field 85
3.6.1. Motion of a Charge in a Constant Uniform Electric Field 85
3.6.2. Motion of a Charge in a Constant Uniform Magnetic Field 86
3.6.3. Motion of a Charge in a Coulomb Field 86
3.7. Fields of Moving Charges. Emission 91
3.7.1. Field of a Uniformly Moving Charge 91
3.7.2. Emission by a Charge Moving with an Acceleration 95
3.7.3. Emission by a Charge Moving Uniformly in a Circle 98
3.8. Electromagnetic Waves 100
3.8.1. Some Properties of Radiation Fields 100
3.8.2. Travelling Waves 100
3.8.3. Emission of Electromagnetic Waves by Oscillating Charges. Energy and Momentum of Waves 102
3.8.4. Free Oscillations of a Field. Standing Waves 104
3.9. Propagation of Light 106
3.9.1. Interference of Electromagnetic Waves 106
3.9.2. Diffraction of Electromagnetic Waves 107
3.9.3. Geometrical Optics 109

CHAPTER 4. ATOMIC PHYSICS AND QUANTUM MECHANICS 110

4.1. Planetary Model of Atom 110
4.2. Experiments on Diffraction of Particles 110
4.3. The Uncertainty Relation 115
4.4. Probability Waves 117
4.4.1. Complex Numbers. Euler's Formula 118
4.4.2. Complex Probability Waves. The Superposition Principle 119
4.4.3. Limiting Transition to Classical Mechanics 121
4.5. Electron in an Atom 123
4.5.1. Energy and Its Quantization 123
4.5.2. Angular Momentum and Its Quantization 128
4.5.3. Probability Amplitudes and Quantum Numbers 130
4.6. Many-Electron Atom 132
4.6.1. Spin of an Electron 132
4.6.2. Systems of Identical Particles. Quantum Statistics 134
4.6.3. Atomic Quantum States 137
4.7. Quantization of Atomic Radiation 139
4.7.1. Quantum Transitions. Line Spectra 139
4.7.2. Photon. The Concept of Parity. Selection Rules 140

4.8. Photon-Electron Interaction. The Photoelectric Effect. The Compton Effect 146
4.9. Simultaneous Measurement of Quantities and the Concept of the Complete Set of Measurable Quantities 150
4.10. Molecules 151

CHAPTER 5. MACROSCOPIC BODIES AS AGGREGATES OF PARTICLES. THERMAL PHENOMENA 155

5.1. The Basic Problem of Statistical Physics 155
5.2. Macroscopic Quantities. Fluctuations 157
5.3. Statistical Analysis of the Gas Model 159
5.3.1. Computer Experiments 159
5.3.2. Reversibility of Microscopic Processes in Time and Irreversibility of Macroscopic Processes 160
5.4. Entropy 161
5.5. Temperature 162
5.6. Equilibrium Distribution of Particles in a Body 167
5.7. Thermodynamic Relations 172
5.8. Ideal Gas 176
5.8.1. Matter and Its States 176
5.8.2. Classical and Quantum Ideal Gases 176
5.8.3. Equation of State for an Ideal Gas 178
5.8.4. Heat Capacity of an Ideal Gas 181
5.8.5. Reversible Thermal Processes 184
5.9. Statistics and Thermodynamics of Radiation 188
5.10. Crystals 194
5.10.1. Crystal Lattice 194
5.10.2. Types of Lattice Bonds 195
5.10.3. Mechanical Properties of Crystals 196
5.10.4. Electron Energy Spectra of Crystals 204
5.10.5. Lattice Heat Capacity 206
5.10.6. Electron Gas in Metals 213
5.11. Phase Transitions 218

CHAPTER 6. MACROSCOPIC MOTION OF MEDIA. NONEQUILIBRIUM PROCESSES 225

6.1. Nonequilibrium States of Bodies 225
6.2. Macroscopic Motion 226
6.3. Equations of Hydrodynamics of an Ideal Liquid 228
6.3.1. Matter Conservation Law in Hydrodynamics 228
6.3.2. Equation of Motion in Hydrodynamics 231
6.4. Hydrodynamic Analysis of Problems on Viscous Flow, Heat Conduction, and Diffusion 233
6.4.1. Viscosity 233
6.4.2. Flow of a Viscous Liquid Through a Tube 235
6.4.3. Heat Conduction 237

6.4.4. Heat Transfer Between Two Walls 238
6.4.5. Diffusion. Dissolution of a Solid in a Liquid 240
6.5. Kinetic Coefficients in Gases and Their Connection with the Molecular Parameters 242
6.5.1. The Concept of Mean Free Path of Molecules 243
6.5.2. Molecular Treatment of the Diffusion Process 246
6.5.3. Diffusion as a Random Motion of Particles 248
6.5.4. Relations Between Kinetic Coefficients 251
6.6. Resistance to the Motion of Solids in a Liquid 252
6.6.1. Similitude Method. The Reynolds Number 252
6.6.2. Drag at Low Velocities 254
6.6.3. Drag at High (Subsonic) Velocities 257
6.7. Instabilities in Hydrodynamics 259
6.7.1. Transition from Laminar to Turbulent Flows 259
6.7.2. Boundary Layer 260
6.7.3. Turbulent Viscosity and Thermal Diffusivity 262
6.7.4. Transition from Molecular to Convective Heat Transfer. Solar Granulation 263
6.8. Oscillations and Waves in a Liquid 266
6.8.1. Various Forms of Wave Motion 266
6.8.2. Wave Characteristics 267
6.8.3. Linear and Nonlinear Waves 269
6.8.4. Solitons and Other Nonlinear Effects 269
6.8.5. Highly Perturbed Media 270
6.8.6. Oscillations of a Charged Drop and the Fission of Heavy Nuclei 271
6.9. Macroscopic Motion of Compressible Media 274
6.9.1. Generalized Form of the Bernoulli Equation 274
6.9.2. Compressibility Criterion for a Medium and the Velocity of Sound 275
6.9.3. Flow in a Tube with a Varying Cross Section 276
6.9.4. Laval Nozzle 277
6.10. Shock Waves 278
6.10.1. Propagation of Perturbations in a Compressible Gas Flow 278
6.10.2. General Relations for a Shock Wave 281
6.10.3. Shock Waves in an Ideal Gas 285
6.10.4. The Problem on a High-Intensity Explosion in the Atmosphere 289
6.11. Hydrodynamic Cumulative Effects 290
6.11.1. Cumulative Jets 291
6.11.2. Bubble Collapse in a Liquid 296

6.11.3. Converging Spherical and Cylindrical Shock Waves 297
6.11.4. The Role of Instabilities in Suppressing Cumulation 297
6.11.5. Emergence of a Shock Wave on the Surface of a Star 298
6.12. Cavitation in a Liquid 299
6.13. Highly Rarefied Gases 301
6.14. Macroscopic Quantum Effects in a Liquid 304
6.15. Generalizations of Hydrodynamics 307

CHAPTER 7. ELECTROMAGNETIC FIELDS IN MEDIA. ELECTRICAL, MAGNETIC, AND OPTICAL PROPERTIES OF SUBSTANCES 309

7.1. Superconductivity 309
7.2. Electrical Conductivity of Metals 310
7.3. Direct Current 315
7.4. Dielectric Conductance 319
7.4.1. Electrons and Holes. Exciton States 319
7.4.2. Semiconductors 320
7.5. Electric Fields in Matter 322
7.5.1. Field Fluctuations in a Substance 322
7.5.2. Electrostatic Fields in Metals 324
7.5.3. Electrostatic Fields in Insulators. Polarization of a Substance 325
7.6. A Substance in a Magnetic Field 330
7.6.1. Diamagnetic Effect 330
7.6.2. Paramagnets. Orientation Magnetization 333
7.6.3. Spontaneous Magnetization. Ferromagnetism 335
7.6.4. Magnetic Properties of Superconductors. Quantization of Large-Scale Magnetic Flux 338
7.7. Alternating Currents and Electromagnetic Waves in a Medium. Optical Properties of Media 342
7.7.1. A.C. Fields and a Substance 342
7.7.2. Induced EMF 343
7.7.3. A.C. Circuits. Solutions of Differential Equations 344
7.7.4. Generation of Electromagnetic Waves 353
7.7.5. Some Laws of Optics and the Velocity of Propagation of Electromagnetic Waves in a Medium. Reflection and Refraction of Waves 355
7.7.6. Refractive Index of Insulators. Dispersion and Absorption of Light 361
7.7.7. Refractive Index of Metals. Skin Effect. Transparency of Metals to Hard Radiation 364
7.7.8. Nonlinear Optics Effects 365
7.7.9. Lasers 369

CHAPTER 8. **PLASMA** 373

8.1. General Remarks 373
8.2. Quantum Effects in Plasma. Tunneling of Nuclei Through a Potential Barrier 374
8.3. Relativistic Effects in Plasma. Mass Defect in Nuclear Fusion and Energy Liberated in the Process 379
8.4. Plasma Statistics. Equation of State for Plasma. Thermal Radiation of Plasma 380
8.5. Plasma Kinetics. Mobility of Ions and Its Relation with Diffusion. Electrical Conductivity of Plasma 384
8.6. Magnetohydrodynamics and Plasma Instabilities. Tokamaks 385
8.7. Oscillations and Waves in Plasma. Propagation of Radio Waves in the Ionosphere 388

CHAPTER 9. **STELLAR AND PRESTELLAR STATES OF MATTER** 392

9.1. State of Matter at Ultrahigh Temperatures and Densities 392
9.2. Stars as Gaseous Spheres 394
9.2.1. Calculation of Pressure and Temperature at the Centre of a Star 394
9.2.2. Temperature of the Surface and the Total Emissive Power of a Star 396
9.2.3. Energy Transfer in Stars 396
9.3. Sources of Stellar Energy 397
9.3.1. Analysis of Possible Sources of Stellar Energy 397
9.3.2. Nuclear Reactions of the Proton-Proton Cycle 399
9.4. White Dwarfs 401
9.4.1. Possible Evolution of Stars of the Type of the Sun 401
9.4.2. Density and Size of White Dwarfs 401
9.4.3. Limiting Mass of White Dwarfs 403
9.5. Superdense Neutron Stars 404
9.5.1. Size of Neutron Stars 404
9.5.2. Rotation and Magnetic Fields of Neutron Stars 405
9.5.3. Radio Emission by Pulsars 406
9.5.4. Internal Structure of Neutron Stars 408
9.5.5. Gravitational Effects in the Vicinity of a Neutron Star 409
9.6. Gravitation and Relativity 411
9.6.1. Equivalence Principle 411
9.6.2. Geometry and Time in Noninertial Reference Frames 412
9.6.3. Einstein's Equations 413

9.7. Expansion of the Universe 413
9.7.1. Friedman's Cosmological Solutions 413
9.7.2. Discovery of "Expansion" of the Universe 415
9.7.3. Critical Density 416
9.8. Hot Universe 418
9.8.1. Discovery of Background Thermal Radiation 418
9.8.2. Charge-Asymmetric Model of Early Universe 419
9.8.3. Change in Density and Temperature of Prestellar Matter in the Process of Cosmological Expansion 421
9.8.4. State of Aggregation at Early Stages of Evolution of Hot Universe 422
9.9. Fusion of Elements in Stars 425
Concluding Remarks 432
Appendices 433
Subject Index 447

Chapter 1
UNITY OF NATURE

1.1. Hierarchy of Natural Objects

In the context of the present discussion, the term "hierarchy" stands for a ladder of objects having different properties or different degrees of complexity. The following hierarchy of objects can be observed in the world surrounding us (this has become possible following many centuries of the study of nature by man): *elementary particles-nuclei-atoms-molecules-macroscopic bodies* (*crystals*, *liquids*, *gases*, *plasma*)-*planets-stars-galaxies-the Universe*.

Man usually interacts with macroscopic bodies, and is himself such an object. Man as a researcher always strives to explore nature in its two extreme manifestations, viz. the smallest (microcosm) and the largest (macrocosm). Each link in the chain *macroscopic body-molecule-atom-nucleus-elementary particle* signifies a landmark in the quest for knowledge. It is noteworthy that if we try to follow this chain in the reverse direction from a particle to a macroscopic body (assuming the properties of particles as given and trying to determine through them the properties of aggregates of particles, i.e. of the macroscopic bodies), the problem turns out to be quite complicated. Thus, even today there is no consistent microscopic theory describing the liquid state of matter.

Let us now consider the entire hierarchy of the basic objects in nature and characterize it briefly.

1.1.1. Elementary Particles. Elementary particles are the simplest basic structure known to date. This, however, does not mean that they have simple properties. The behaviour of elementary particles is described with the help of profound physical theories combining the theory of relativity and the quantum theory.

All known elementary particles[1] are divided into two

[1] With the exception of photons.

groups, viz. **hadrons** and **leptons**. In its turn, the group of hadrons consists of **baryons** and **mesons**:

Hadrons → Baryons (proton (p), neutron (n), hyperons (Λ, Σ, Ξ, Ω)[2]).
Hadrons → Mesons (π-mesons, K-mesons).

Leptons can be represented in the form:

Leptons → electron (e) and electron neutrino (ν_e),
Leptons → muon (μ) and muon neutrino (ν_μ),
Leptons → tau-lepton (τ) and tau-neutrino (ν_τ).

Moreover, the group of hadrons contains two large families of **resonances**, viz. **baryon resonances** and **meson resonances**. Over 300 resonances have been discovered so far.

The last remark may well intrigue the reader. Is the total number of particles not too large? Are they all elementary particles indeed? These questions confounded and continue to confound physicists even today. It was established in the sixties that hadrons may be grouped together into certain families of particles having nearly identical properties. The properties of such particles and the number of particles in each group could be determined by assuming that hardrons have a composite structure, i.e. consist of truly elementary particles called **quarks**. According to this hypothesis, baryons are composed of three quarks (antibaryons are composed of three antiquarks), while mesons are formed by a quark and an antiquark. All hadrons known at that time could be built by postulating the existence of only three types of quarks.

In the beginning, when the quark model of hadrons was constructed, quarks were treated as purely mathematical structural elements which provide a very convenient representation of hadrons. However, subsequent experiments on the scattering of high-energy electrons by nucleons revealed the existence of point-like charged formations inside the nucle-

[1] Greek letters are used for hyperons.

ons. Naturally, these formations were identified as quarks. Quarks have not been detected in free state so far, and theoretical considerations indicate that quarks cannot exist in free state.

Subsequent discoveries in elementary-particle physics necessitated the introduction of two new types of quarks. In all, six different types of quarks are assumed to exist.

Like quarks, leptons are also considered to be structureless point particles. They are also represented by a family of six members.

Quarks and leptons can claim to be truly elementary particles at present.

Gluons, photons, and **heavy intermediate bosons** are particles which serve as carriers of interaction. At present, all interacting elementary particles are presented as a sort of ball game: quarks interact by throwing gluons at one another, photons are exchanged in interactions between electrically charged particles, while heavy intermediate bosons are responsible for slow decay of particles and for extremely weak interaction of all types of neutrinos with matter. In view of the large mass of intermediate bosons, the range over which they are exchanged is very small ($\leqslant 10^{-18}$ m). Hence the process of exchange of heavy intermediate bosons will become possible only when the interacting particles approach each other to very small distances, which is a very rare event (on the time scale used in the microcosm).

We shall describe the classification of interactions at a later stage. For the present, let us consider some properties of elementary particles. Protons, electrons, photons, and (apparently) all types of neutrinos are **stable particles**, i.e. can exist in free state for an indefinitely long time. Resonances, having a lifetime of about 10^{-23} s, are the shortest-lived formations. The record of longevity in the microcosm is held by the neutron in the free state, having a lifetime of about 10^{3} s. Hyperons live for about 10^{-10} s after which they disintegrate. Mesons are characterized by a lifetime of about 10^{-8} s. As regards leptons, the muon has a lifetime of 10^{-6} s, while the tau-lepton lives for less than 10^{-12} s.

The world of short-lived elementary particles can be created artificially with the help of **accelerators** which are huge engineering installations. High-intensity beams of protons or electrons are produced in accelerators and are accelerated

to high velocities with the help of electric fields. As a result of collisions with the matter in the target (for example, liquid hydrogen in the bubble chamber which also serves as the detector), high-energy protons or electrons produce new short-lived elementary particles. As the energy of colliding particles increases, new levels in their structure are revealed. High-energy particle beams are used to "feel" the submicroscopic structure of the investigated objects. This necessitates the construction of particle accelerators with higher and higher energies.

Short-lived elementary formations in nature may participate in the most complicated interactions between matter and field, e.g. at the "initial" stages of evolution of the Universe, in the formation of astrophysical objects like black holes, in the formation of the core of neutron stars.

The masses of stable elementary particles are distributed as follows. The proton has a mass $m_p \simeq 1.67 \times 10^{-27}$ kg, while the electron mass is $m_e \simeq 0.91 \times 10^{-30}$ kg. Photons, gluons, and (probably) all types of neutrinos have a zero rest mass; in the latest experiments, the value of the rest mass of the electron neutrino was found to be $\leqslant 10^{-35}$ kg. The hyperon mass is larger than the proton mass, while the meson mass is smaller than the proton mass.

The elementary charge is defined as the charge with absolute magnitude $e = 1.6 \times 10^{-19}$ C. Almost all elementary particles have a charge equal to $+e$ or $-e$, or are electrically neutral. The only exception to this are certain resonances having multiple electron charge.

Particles with a fractional charge have not been observed. Quarks, which are truly elementary particles, have a charge that is multiple of $e/3$. However, quarks cannot exist outside their "cradle", viz. the heavy particle formed by them.

A combination of relativistic and quantum concepts, formulated mainly in the thirties, led to one of the most outstanding predictions in physics, viz. the existence of antiparticles. A particle and its corresponding antiparticle have the same lifetime, mass, and equal but opposite charge. Moreover, particles and antiparticles have a whole lot of specific "intrinsic" properties which differ only in sign. The most striking property of particle-antiparticle pairs is the ability of such particles to *annihilate* (destroy) each other and be converted into other particles.

Antiparticles of almost all known elementary particles have experimentally been discovered by now. It is interesting to note that some particles are identical to their antiparticles. Examples of such particles are the photon and π^0-meson.

Antiparticles may combine to form **antimatter**. For example, antihelium-3 was obtained at the accelerator at the Institute of High-Energy Physics in Serpukhov. The nucleus of an antihelium-3 atom consists of two antiprotons and one antineutron. This nucleus is surrounded by a cloud of **positrons** (antielectrons). The neutron and antineutron, which are electrically neutral particles, have equal but opposite magnetic moments (relative to the intrinsic mechanical moment).

In spite of the microscopic symmetry between particles and antiparticles, there is no evidence so far of the existence in the Universe of regions containing any perceptible amount of antimatter.

It should be observed that particles and corresponding antiparticles interact with the gravitational field in an identical manner, thus indicating the absence of "antigravity" in nature.

1.1.2. Nuclei. Atomic nuclei are bound systems of protons and neutrons[3]. *The mass of a nucleus is always slightly lower than the sum of the masses of free nucleons constituting the nucleus.* This is a relativistic effect determining the binding energy of the nucleus.

Unlike mass, *the charge of a nucleus is exactly equal to the sum of the charges of protons constituting it.* Nuclei with a charge from $1e$ to $107e$ and with the number of nucleons from 1 to about 260 are known at present. Nuclei having 2, 8, 20, 28, 50, 82 or 126 protons or neutrons are especially stable, i.e. have the highest binding energy. Hence these numbers are called **magic numbers**.

The number density of particles in a multinucleon nucleus is of the order of 10^{44} nucleons/m^3, while the mass density is about 10^{17} kg/m^3. The "nuclear radii" vary from 2×10^{-15} m (helium nucleus) to 7×10^{-15} m (uranium nu-

[3] Protons and neutrons are often referred to by a single term *nucleon* meaning a nuclear particle.

cleus). Nuclei have the form of a prolate or oblate ellipsoid, or a more complex shape.

As a quantum-mechanical system, a nucleus may exist in different discrete excited states. In the ground state, nuclei may be **stable** or **unstable**. The time during which half the number of nuclei decay out of any macroscopic amount of unstable nuclei is called the **half-life.**[4] The half-lives of known elements vary from about 10^{18} years to 10^{-10} s.

It was mentioned above that free neutrons are unstable particles decaying into a proton p, an electron e^-, and an electron antineutrino $\tilde{\nu}_e$:

$$n \to p + e^- + \tilde{\nu}_e$$

(it should be remarked that the neutron has a slightly larger mass than the proton). However, neutrons constituting stable atomic nuclei are stable. In unstable nuclei, neutrons may undergo the above transformation. As an example, let us consider the decay of a boron nucleus which is transformed into a carbon nucleus:

$$^{12}_{5}\mathrm{B} \to {}^{12}_{6}\mathrm{C} + e^- + \tilde{\nu}_e.$$

Astrophysical objects include neutron stars which are sort of huge atomic nuclei composed mainly of neutrons. The mechanical equilibrium of neutron stars is maintained by gravitational forces. It is still not clear whether multineutron nuclei can exist in the microcosm only on the basis of nuclear interaction.

1.1.3. Atoms and Molecules. Nuclei have a positive electric charge and are surrounded by a cluster of negatively charged electrons. Such an electrically neutral formation is called an **atom.** *An atom is the smallest structural unit of chemical elements.*

Unlike the "dense packing" of nucleons, atomic electrons form loose and openwork **shells.** The occupation of atomic levels around the nucleus takes place according to strict rules. The electrons on the highest levels of the atomic build-

---

[4] We are speaking of statistical mean time since nuclear decay follows probability laws.

ing determine the reactivity of atoms, i.e. their ability to enter into combination with other atoms. We are trespassing into the area of chemistry here, and the arbitrariness in setting a boundary between physics and chemistry becomes quite obvious in this case.

Not all atoms are capable of combining with one another. A combination is possible only if the joint "top floor" is completely filled by electrons. Such a combination is called a **molecule**. *A molecule is the smallest structural unit (building block) of a complex chemical compound.* The number of possible combinations of atoms determining the number of chemical compounds is as large as 10^6.

Some atoms (for example, carbon and hydrogen) are capable of forming complex **molecular chains**[5] which serve as the basis for the creation of more complicated structures (macromolecules) exhibiting biological properties.

1.1.4. Macroscopic Bodies. Under certain conditions, atoms or molecules of the same type may combine to form giant aggregates called **macroscopic bodies (matter)**. Simple matter is atomic in nature, while complex matter is molecular. For the sake of simplicity, we shall confine ourselves to atomic bodies.

At very low temperatures, all objects are **crystals**.[6] The mutual arrangement of atoms in a crystalline body is regular. In such bodies, atoms cannot move far away from their equilibrium positions called the **crystal lattice sites**. Their motion is mainly confined to oscillations about the lattice sites.

The geometry of the crystalline state of matter under usual temperatures and pressures is distinguished by an extraordinary variety in spite of the fact that there is a limited number of crystal lattices. The properties of matter are determined not only by the nature of atoms but also by their mutual arrangement. An example of this is the case of diamond and graphite, both of which are formed by the same carbon atoms but have different crystal lattices.

---

[5] Over 3×10^6 organic compounds of carbon had been discovered by the seventies, while the compounds of all the remaining elements did not exceed 10^5 in number.

[6] The only exceptions are ^{4}He and ^{3}He, which remain in the liquid state right down to 0 K (under pressures below 25 atm).

Bodies may differ strongly in their mechanical, thermal, electrical, magnetic, and optical properties. Knowing the atomic structure of bodies and the dependence of their properties on it, we can endeavour to form new materials.

From the crystalline state of matter, we go over to the **liquid state**. A considerable increase in temperature results in the crystal-liquid **phase transition (melting)**. Each material has its own definite melting point. In the liquid state, atoms are no longer strictly localized or associated with some definite positions in a body. Thermal motion in the liquid is quite complicated in nature.

Upon a transition from liquid to **gaseous state** (under atmospheric pressure), matter almost completely loses its individuality. This is associated with a low density of gaseous matter. In rarefied gases, there is essentially no mutual influence of atoms, which means that no individual atomic structure is manifested. All substances in gaseous state (under normal conditions) obey the same laws to a considerably high degree of accuracy.

A subsequent considerable increase in temperature (to about 10^4-10^5 K) of the medium results in **ionization** of atoms, i.e. in their decomposition into ions and free electrons. Such a state of matter is called **plasma**.

Since ions and electrons carry uncompensated electric charges unlike atoms, they exert a considerable influence on one another. In contrast to gases, plasma may exhibit collective properties. This makes plasma similar to the **condensed state** of matter, i.e. to solids and liquids. All kinds of elastoelectric oscillations can easily be excited in plasma.

For the sake of reference, we quote the densities of macroscopic bodies under normal conditions: solids and liquids have a density of about 10^{28}-10^{29} atoms/m^3, gases have about 10^{25} molecules/m^3, while the density of particles in artificially created plasma is about 10^{19} m^{-3}.

1.1.5. Planets. The next step in the hierarchy of natural objects concerns macroscopic bodies on astronomical scale, viz. **planets**.

Investigations of the planets of the solar system have actually just begun. Automatic interplanetary stations and landing vehicles have already been sent to Venus and Mars. Man has already set foot on the Moon.

What are the results of these investigations? Firstly, the conditions under which planetary matter exists are considerably different from terrestrial conditions. Secondly, no new chemical elements have been discovered in the planetary matter.

The study of the internal structure of planets is quite a complicated task. Even our knowledge about our "own home", viz. the Earth, is extremely inadequate. The internal pressure in the central part of the Earth is estimated at about 10^7 atm (10^{12} Pa). What type of changes does matter undergo at such pressures?

While matter at moderate pressures is characterized by a large variety and diversity in its properties and exhibits an extremely sharp and nonmonotonic dependence on chemical composition, there is a clear tendency towards a "smoothing" of properties under compression. The latter circumstance can easily be explained. As a matter of fact, the outermost electron shells of atoms, which are responsible for the above-mentioned nonmonotonicity, cease to exist under pressures of the order of 10^7-10^8 atm. This is so because the electrons in these shells are ripped off the atoms and become collectivized. The inner electron shells, however, are compressed, and the distribution of electron density changes rather slowly as we go over from one substance to another.

Physicists have learnt the technique to produce pressures of the order of 10^7 atm in the laboratory by shock compression of bodies. True, such pressures can be maintained only for brief intervals of time. However, investigations are under way to create presses for producing a static pressure of this order of magnitude.

Studies in the field of high-pressure physics have led to the synthesis of a whole range of new materials including artificial diamonds. It is hoped that metallic hydrogen will be obtained at pressures of about 10^7 atm. The properties of metallic hydrogen are likely to be quite different from those of ordinary dielectric crystals of molecular hydrogen, which can be obtained under normal pressures at low temperatures. Theoretical physicists believe that crystals of metallic hydrogen should be superconducting at room temperature. If this assumption is found to be true, we shall witness a revolution in electrical technology.

Jupiter is one of the largest planets in the solar system.

It is about ten times the size of the Earth and apparently exists in fluid state. It is rich in hydrogen, and the central part of this planet is surely made of metallic hydrogen.

1.1.6. Stars. Galaxies. Universe. Strange though it may seem, physicists have a much better understanding of the internal structure of stars than of planets.

The core of the Sun is characterized by a temperature of about 10^7 K and a pressure of about 10^{11} atm. Under these conditions, matter exists in the state of a completely ionized plasma, i.e. bare nuclei and free electrons. These conditions are conducive for fusion reactions involving the fusion of hydrogen nuclei into helium nuclei. This nuclear reaction serves as the source of energy of stars.

As stars lose their energy, they contract. Free electrons are sort of squeezed into the nuclei. Electrons are captured by protons which are thus transformed into neutrons. This is accompanied by the emission of a neutrino. As a result of such a reaction, the nuclear charge is reduced (the mass of the nucleus remains the same), and this generally leads to a decrease in the binding energy of the nucleus. In the long run, nuclei containing too many neutrons become unstable and disintegrate. The matter of stars is a superdense neutron gas.

Neutron stars have been discovered in recent years. They have a mass comparable with that of the Sun but are smaller in size by a factor of 10^5. The density of neutron stars is comparable with that of the nuclear matter, viz. 10^{17} kg/m^3. The mass of 1 cm^3 of such a star is hundreds of million tons!

Galaxies are stellar systems. The number of stars in the galaxies is of the order of 10^9-10^{12}. The mass of a star is of the order of 10^{30} kg (as that of the Sun), while the mass of the Galaxy is of the order of 10^{41} kg.

The part of the **Universe** familiar to us contains about 10^{11} galaxies. The total number of protons and neutrons in the part of the Universe familiar to us is of the order of 10^{80}.

Leaving aside the proofs and the methods of measurement leading to such numerical estimates, the fact remains that we are aware of them. Obviously, some of these numbers are rough estimates.

Astronomical data show that galaxies in the Universe tend to "run away" from one another. The expansion of the Uni-

verse indicates that somewhere in the distant past it occupied a very small volume. In turn, this means that prestellar matter was superdense and extremely hot (having a temperature of the order of 10^{13} K). At such temperatures, matter may consist primarily of radiation, viz. photons and neutrinos.

This "hot" model of the Universe leads to a number of results that can be verified experimentally. The discovery of **relict photons** in recent years (residues of prestellar state of matter) is a sound evidence in favour of our concepts about the earliest stages in the evolution.

1.2. Four Types of Fundamental Interactions

1.2.1. Bound Systems of Objects. Interactions. It was mentioned in the previous section that qualitatively different kinds of bound systems exist in nature. For example, a nucleus is a bound system of protons and neutrons, an atom is a bound system of nuclei and electrons, macroscopic bodies are bound systems of atoms or molecules, and the solar system is a "bunch" of planets and the heavy star.

The existence of bound systems of objects indicates that there must be something tying parts of the system into a single entity. A partial or complete "rupture" of the system requires the expenditure of energy. The mutual interaction of the parts of the system is characterized by the **interaction energy**, or simply **interaction**.

At present, it is believed that all interactions of any type of objects can be reduced to a limited class of fundamental interactions, viz. gravitational, electromagnetic, strong, and weak.

1.2.2. Gravitational Interactions. The attraction between objects and the Earth, the existence of the solar system and stellar systems (galaxies) are due to the **gravitational forces** or, in other words, the gravitational interaction. Gravitational interaction is universal, i.e. can be applied to all micro- and macroobjects. However, it is significant only for giant objects of astronomical masses and for the formation of the structure and evolution of the Universe as a whole. Gravitational interaction attenuates very

rapidly for objects of small masses and is practically insignificant for nuclear and atomic systems.

The manifestations of gravitation was one of the first effects to be investigated quantitatively. This is not surprising, since the source of gravitation is the mass of the objects, and the range of gravitational interaction is unlimited.

1.2.3. Electromagnetic Interactions. These interactions are responsible for the bonds in atoms, molecules, and ordinary macroobjects. The ionization energy of an atom, i.e. the energy of dissociation of an electron from the nucleus, determines the electromagnetic interaction in the atom. The heat of vaporization, i.e. the energy of liquid-vapour transition (at atmospheric pressure), determines, albeit roughly, the value of the intermolecular interactions in a body. These interactions are of electromagnetic origin.

1.2.4. Strong (Nuclear) Interactions. The presence of like-charged protons and neutral particles (neutrons) in the nucleus indicates that there must be interactions much stronger than the electromagnetic interactions since otherwise the formation of the nucleus would be impossible. These interactions, called strong interactions, are manifested only within the nucleus.

In order to determine certain properties of nuclear interactions, let us consider the **mirror nuclei**. These nuclei are obtained from one another through a replacement of protons by neutrons and vice versa. Such nuclei have nearly identical physical properties (binding energy, energy spectrum, etc.). An example of mirror nuclei are the ^{11}B and ^{11}C nuclei. The physical equivalence of mirror nuclei indicates that strong interactions of the type p-p, p-n, and n-n must be identical. Of course, there is a slight difference between these types of interactions due to electromagnetic forces.

Strong interactions take place not only between nuclear particles but also between baryons and mesons.

1.2.5. Weak Interactions. It was mentioned above that among the vast multitude of elementary particles, only p, e, γ, ν_e, ν_μ, and ν_τ are stable. Due to "internal reasons", all unstable free particles are transformed into other parti-

cles in a certain characteristic time. For example, resonances decay in about 10^{-23} s as a result of strong interactions, while neutral π^0-mesons are subjected to electromagnetic interactions and decay in about 10^{-16} s. Slow decays are characterized by a period of 10^{-10}-10^{-6} s. These decays are caused by weak interaction.

A large number of slow decays of elementary particles is accompanied by the emission of a neutrino. This particle has an extremely weak interaction with matter. As it passes through a medium having an ordinary density, the path length between two successive collisions of neutrino with the particles of matter is astronomically large (of the order of 10^{17} km). This means, for example, that the Earth having a radius of 6.4×10^3 km is absolutely "transparent" to the neutrino flux.

How then were the physicists able to not only detect neutrinos but also study their properties? After all three types of neutrinos, viz. ν_e, ν_μ, and ν_τ, have been detected. Low-energy antineutrinos $\tilde{\nu}_e$ with a flux of 10^{17} particles/($m^2 \cdot s$) are produced in a nuclear reactor. Naturally, the probability of detecting a neutrino for such a large value of the flux sharply increases. This was first done in 1953. A flux of high-energy neutrinos (up to 10^{11} eV) can be generated with the help of charged particle accelerators.

The general set-up for neutrino investigations being carried out at the Institute of High-Energy Physics in Serpukhov (USSR), CERN (Geneva), Brookhaven and Batavia (USA) can be described as follows. A beam of protons emerging from an accelerator interacts with the target nuclei, generating π- and K-mesons in the process. These mesons first fall on a focus unit (which selects particles of the required charge) and then in a decay channel (of length varying from 30 m to 1300 m depending on the accelerator energy) where a beam of neutrinos is formed during decay reactions involving the creation of muons and neutrinos. The decay channel terminates in a massive steel absorber or muon filter. The flux of neutrinos passing through the filter without getting absorbed falls on the detector (target) in the form of a bubble chamber of volume up to 20 m^3 or a spark chamber and scintillation counters containing layers of Fe or Al filters. Detectors have cross sections of 3-4 m, are up to 16 m in length, and weigh up to 1000 tons.

Neutrino experiments on accelerators not only help in studying the properties of neutrinos but are also used for investigating weak interactions and properties and structure of other elementary particles.

1.2.6. Comparative Estimates for the Intensity of All Types of Interactions. If we consider only elementary particles, the intensity of various interactions with respect to strong interactions is distributed as follows:

Strong	~ 1
Electromagnetic	$\sim 10^{-3}$
Weak	$\sim 10^{-14}$
Gravitational	$\sim 10^{-40}$

In questions concerning the structure and evolution of the world as a whole, gravitation plays a decisive role. However, the investigations of specific astrophysical objects (stars, pulsars, quasars, etc.) are impossible without a consideration of all fundamental interactions.

1.2.7. Fields and Matter. The aggregate of elementary particles and their interactions is manifested macroscopically in the form of matter and field.

Unlike matter, field has peculiar properties. The physical reality of electromagnetic field is evident at least from the fact of the existence of radio waves having a finite velocity of propagation. For example, a delay of about 2.5 s was observed in the radioecho while determining the position of the Moon with the help of radio waves.

Moving charged particles are the source of electromagnetic field. The interaction of charges takes place through the mechanism particle-field-particle. Field is the carrier of interaction. Under certain conditions, the field may detach itself from the source and propagate freely in space. Such a field has wave properties.

Let us consider the source of information on the state of stellar matter. Atomic processes taking place in the outer layers of stars are accompanied by the emission of electromagnetic waves. One such process is the excitation of atoms leading to the emission of a series of specific "portions" of electromagnetic field (spectrum). Each chemical element

has its own emission spectrum. For example, an analysis of sunlight (light is nothing but electromagnetic radiation) with the help of optical instruments reveals the chemical composition and concentration (in %) of elements in the outer layers of the Sun.

Nuclear processes occurring in the bowels of stars are accompanied by the emission of the all-pervading neutrinos. Hence neutrinos must provide information about the type of fusion reactions taking place inside the Sun.

1.3. Space and Time

1.3.1. Scales of Space and Time in Nature. The phenomena and processes involving interacting objects occur in space and time. Space and time serve as the "arena" in which events take place. If an event is characterized by its place of occurrence and the instant at which it happens, the space-time relations impose certain constraints on the possible course of events.

The hierarchy of objects and interactions described above can be assigned spatial and time characteristics. For example, the upper boundary of the region in which strong and weak interactions take place has a size $\sim 10^{-15}$ m. The electromagnetic and gravitational interactions have an unlimited range, and this explains the possibility of macroscopic manifestation of these interactions.

Interactions differ significantly in their time characteristics. Processes occurring under the influence of strong interactions are characterized by time durations of $\sim 10^{-23}$ s. For processes associated with electromagnetic interactions of particles, the characteristic time is $\sim 10^{-15}$ s. Finally, for processes whose course is dictated by weak interactions, the characteristic time is $\geqslant 10^{-8}$ s.

Bound systems of protons and neutrons (nuclei) are characterized by a spatial region of about 10^{-15} m. The characteristic region occupied by electrons in an atom is $\approx 10^{-10}$ m.

Let us consider the space-time scale of astrophysical objects. The mean radius of the Earth is $\simeq 6.4 \times 10^6$ m, and its lifetime is $\simeq 4.6 \times 10^9$ years. The radius of the Sun is 7×10^8 m, while its age is $(5\text{-}10) \times 10^9$ years. Galaxies have a size of about 10^{21} m, and the age of the stellar matter

is (5-10) $\times$ 10^9 years. The radius of the part of the Universe that we know is $\sim 10^{26}$ m, and its age is $\sim 10^{10}$ years.

1.3.2. Homogeneity of Space and Time. Space and time have definite properties and this naturally affects the course of physical events. The most important of these properties is **homogeneity**. The **homogeneity of space** means that all points in space are physically equivalent, i.e. *a transportation of any object in space does not affect in any way the processes taking place in this object*. Thus, we are sure that atoms have the same properties on the Earth, Moon or Sun.

The **homogeneity of time** must be understood as the physical indistinguishability of all instants of time for free objects. We believe that the regularities observed today in the behaviour of atoms existed many million years ago as well.

Let us illustrate this point by considering some more examples. The same experiment carried out in Moscow and New York gives the same results. This can be treated as the equivalence of different points in space under terrestrial conditions. The laws of buoyancy discovered by Archimedes many centuries ago can be reproduced even today by creating appropriate conditions for their observation. In other words, all instants of time in this case are physically equivalent.

If these seemingly obvious properties of homogeneity of space and time had not existed, it would be meaningless to carry out scientific investigations. Indeed, the absence of homogeneity of space would mean that the laws of physics would be different in Moscow from those in Paris or in New York. The lack of homogeneity of time would mean the absence of advancement of knowledge. The law of floating bodies discovered yesterday would no longer be valid today and investigations would have to be carried out again. In turn, the law discovered today would not be valid tomorrow.

1.3.3. Free Bodies and Inertial Motion. Our space is "plane" in the sense that it obeys all axioms of Euclidean geometry. This fact has been proved experimentally. How would free bodies behave in such a space?

By a **free body**, *we shall mean a body which is so far removed from all surrounding objects that their interaction can be neglected.* Owing to the homogeneity of space, any position of

such a free body can in no way affect its state. The point of space where a body is located is immaterial since the "external conditions" remain the same. If, in addition, we take into account the homogeneity of time as well, i.e. the physical equivalence of all moments of time for a free body, we arrive at an interesting conclusion: the body will *move*, i.e. *change its position regularly with the passage of time.* Moreover, in view of the homogeneity of space and time, the motion will be **uniform**, i.e. *the body will cover equal distances in equal intervals of time.* The motion will also be **rectilinear** in view of the fact that the space is "plane". Such a motion of free bodies is called **inertial motion.**

It should be noted that the properties of space and time on the scale of the Universe or in the vicinity of huge objects having astronomically large masses deviate from the properties of ordinary Euclidean geometry. Here, the space has a peculiar "warping". Inertial motion in such a non-Euclidean space is not rectilinear or uniform.

In the investigation of elementary particles, nuclei, atoms, molecules, and macroscopic bodies, we assume that the space is Euclidean. This is confirmed by experimental studies to a very high degree of accuracy. As an illustration of the inertial motion of bodies, let us consider a rarefied gas. Individual particles in such a gas can be treated as free bodies since the molecular interaction is extremely small and direct collisions of molecules are extremely rare. Molecules perform inertial motion between collisions.

The inertial motion of bodies is the manifestation of specific symmetry of space and time, viz. their homogeneity.

1.3.4. Inertial Reference Frames. The Relativity Principle. In order to study quantitatively the motion of any object, we must have a **reference frame.** By a reference frame we mean a **coordinate system** and a **clock** fixed to the reference body.

For the coordinate system, we shall normally use the **rectangular Cartesian system.** Speaking of clocks, we mean not only the mechanism invented by man but rather any periodic process in nature. For example, we can choose in place of a clock

rotation of the Earth about its axis,
revolution of the Earth in an orbit around the Sun, or
periodic motion of electrons in atoms (atomic clock).

If a coordinate system and a clock are associated with an arbitrarily moving body, the physical phenomena relative to such a reference frame will appear quite complicated even in the simplest case. The quantitative description of phenomena becomes especially simplified if freely moving bodies are chosen as the reference body. Such reference frames are called **inertial.**

All freely moving objects move uniformly along a straight line in inertial reference frames. Any number of reference frames can be chosen, and all of them will move inertially relative to one another.

There are no criteria for preferring one inertial frame to another inertial frame. All inertial reference frames are physically equivalent, and this is confirmed experimentally.

All physical phenomena retain their identity in all inertial reference frames. This means that *the mathematical formulation of the laws of nature must be such that it does not change upon a transition from one inertial reference frame to another.* This statement in physics is called the **relativity principle.**

Owing to the relativity principle, physical laws possess a special degree of perfection associated with their symmetry in the choice of an inertial reference frame. The requirement of this symmetry plays a significant heuristic role in the quest for quantitative relations describing phenomena in new fields of investigations.

Chapter 2
MECHANICS OF A MATERIAL PARTICLE

2.1. Coordinates, Velocity, Acceleration

The mechanical motion of bodies, which is the simplest of all physical phenomena in nature, plays a very important role in physics as a whole. The laws of such a motion include not only the motion of free bodies considered in Chap. 1 but also the motion of interacting bodies. This general type of motion is called **mechanical motion.**

Mechanical motion can conveniently be studied by considering an idealized object, viz. a **material particle**, which means a body whose size can be neglected while describing its motion.[1] To know the motion of a material particle is to know its position in space at all instants of time.

Experience shows that three **coordinates** must be specified for completely defining the position of a material particle in space with respect to any reference body. We fix a certain Cartesian coordinate system with the reference body. The position of a material particle in space can then be characterized by its **radius vector** $\mathbf{r}$ whose components along the coordinate axes are equal to the Cartesian coordinates x, y, and z of the particle (Fig. 2.1).

As a material particle moves, its position in space changes. Accordingly, the radius vector of a material particle can be treated as a function of time t:

$$\mathbf{r} = \mathbf{r}(t). \tag{2.1}$$

The vector function (2.1) is equivalent to three scalar functions $x(t)$, $y(t)$, and $z(t)$.

Mechanical motion may be distinguished by a rapid variation of coordinates with time. Therefore, we introduce the concept of **velocity** $\mathbf{v}$. This quantity is defined as the ratio

---

[1] The material particle approximation is also applicable when the motion of a finite-sized body is known to be translatory.

of the limiting value of the increment of radius vector to the interval of time in which this increment occurs under the condition that this time interval tends to zero.[2] Taking this formulation into consideration, we can write the following expression for the velocity:

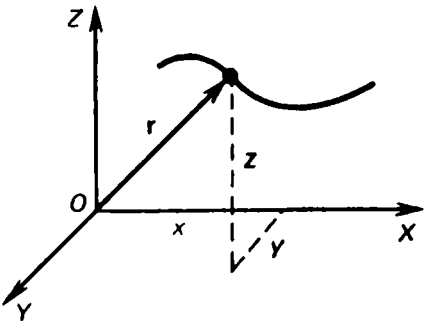


FIG. 2.1

$$\mathbf{v} = \frac{\Delta \mathbf{r}}{\Delta t}. \qquad (2.2)$$

For components of the velocity, say, along the X-axis, we obtain

$$v_x = \frac{\Delta x}{\Delta t}. \qquad (2.3)$$

A material particle may move with a constant or variable velocity. Motion with a constant velocity is called **uniform rectilinear motion**. If the velocity changes, the quantity $\Delta \mathbf{v}/\Delta t$, called the **acceleration**, is nonzero, and the motion is called **accelerated motion.**

2.2. Galilean Transformations

It was mentioned above that it is impossible to study the motion of a body without associating it with some reference frame. It was agreed that we shall use only the inertial reference frames. If the motion $\mathbf{r}(t)$ of a material particle in an inertial reference frame is known, we can find its motion $\mathbf{r}'(t')$ in another inertial frame. For this purpose, we must determine the space-time transformations of quantities as a result of such a transition. This is quite significant since we are speaking of such important concepts as space and time.

---

[2] Mathematically, it can be expressed as follows:

$$\mathbf{v} = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{r}}{\Delta t} \equiv \frac{d\mathbf{r}}{dt}.$$

In other words, the velocity is the derivative of the radius vector with respect to time. For the time being, we shall not introduce the concept of derivative in the text and confine ourselves just to finite differences.

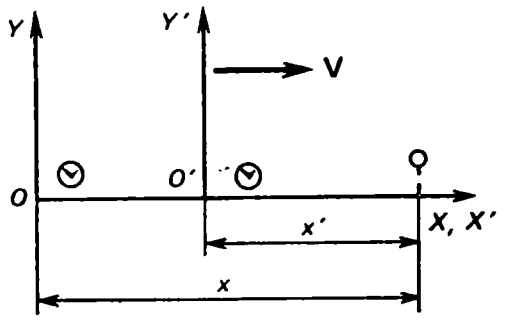


FIG. 2.2

Let us consider two inertial reference frames (X, Y) and (X', Y') (Fig. 2.2). Suppose that at the initial instant of time $t = 0$ the origins O and O' of these frames coincide. Let the reference frame (X', Y') move relative to the reference frame (X, Y) with a constant velocity V along the X-axis. Then the coordinates of the same point in these two reference frames are connected through the relations

$$\left.\begin{aligned} x &= x' + Vt, \\ t &= t'. \end{aligned}\right\} \qquad (2.2.1)$$

The y- and z-coordinates remain unchanged ($y = y'$ and $z = z'$). These transformations are called the **Galilean transformations.**

While the first equation in (2.2.1) is quite obvious, the second requires an explanation. This equation means that time is absolute, i.e. independent of the choice of inertial reference frame. This assumption is experimentally verified for a wide range of physical phenomena.

2.2.1. Absolute Nature of Dimensions and Time Intervals. Let $\Delta x = x_2 - x_1$ be the distance between two points in the reference frame (X, Y). We shall find the relation between Δx and $\Delta x' = x_2' - x_1'$. It follows from (2.2.1) that

$$x_1 = x_1' + Vt,\ x_2 = x_2' + Vt.$$

This means that $x_2 - x_1 = x_2' - x_1'$, or, in other words,

$$\left.\begin{aligned} \Delta x &= \Delta x', \\ \Delta t &= \Delta t'. \end{aligned}\right\} \qquad (2.2.2)$$

These equations express the independence of lengths and time intervals of the choice of inertial reference frames. It can be stated that *the dimensions of a body and the passage of time do not depend on the state of motion of the body.*

2.2.2. Relative Nature of Velocities and the Law of Their Transformation. From Eq. (2.2.1), we have

$$\Delta x = \Delta x' + V \Delta t,$$

or

$$\frac{\Delta x}{\Delta t} = \frac{\Delta x'}{\Delta t} + V \frac{\Delta t}{\Delta t},$$

i.e.

$$\mathbf{v} = \mathbf{v}' + V. \tag{2.2.3}$$

Here, we have used definition (2.1.3) for velocity, as well as the assumption that the velocity of a material particle is directed along the X- and X'-axes. The vector relation (2.2.3) is true for the general case. Here, $\mathbf{v}'$ is the velocity of the material particle in the inertial reference frame (X', Y'), $\mathbf{v}$ is the velocity of the same particle in the reference frame (X, Y), and, as was mentioned above, $\mathbf{V}$ is the velocity of transport of the primed frame relative to the unprimed frame. Thus, *velocity is a relative concept, and its value depends on the choice of the reference frame.* For example, if a body is at rest in some reference frame, it generally moves with a constant velocity relative to all other reference frames. This can be seen clearly from Eq. (2.2.3): if we put $\mathbf{v}' = 0$, then $\mathbf{v}$ becomes equal to $\mathbf{V}$.

2.2.3. Absolute Nature of Accelerations. Let us apply the operator Δ (the difference of two quantities with nearly equal values) to the left- and right-hand sides of Eq. (2.2.3) and consider that $\mathbf{V}$ is a constant vector. This gives

$$\Delta \mathbf{v} = \Delta \mathbf{v}',$$

or

$$\frac{\Delta \mathbf{v}}{\Delta t} = \frac{\Delta \mathbf{v}'}{\Delta t'}. \tag{2.2.4}$$

This relation indicates that *the acceleration of bodies in all inertial reference frames is the same.*

2.3. Law of Motion in Mechanics

It was mentioned above that a free body has a uniform rectilinear motion, i.e. a constant velocity. If, however, the body is not free and interacts with the surroundings, its

velocity will generally not remain constant, and the body will experience an acceleration. Thus, the acceleration can be expressed in the form

$$\frac{\Delta \mathbf{v}}{\Delta t} \propto \mathbf{F},$$

where **F** is the force that serves as a measure of the external influence on the body.

The above relation cannot be considered complete since it does not contain the mechanical properties of the body in any form. Such a property of the body is its **mass** m. In the final form, we obtain the relation

$$m \frac{\Delta \mathbf{v}}{\Delta t} = \mathbf{F}. \tag{2.3.1}$$

This relation expresses the **law of motion in mechanics.**

This mechanics is called **Newtonian**, or **classical**. It has definite limits of applicability, which will be considered below.

Like any other fundamental law, the law of motion (2.3.1) in mechanics must satisfy the relativity principle, i.e. the mathematical form of the law must not depend on the choice of inertial reference frames. For Newtonian mechanics, this means that *the equation of motion* (2.3.1) *for a material particle does not change its form under Galilean transformations.*

Before applying transformations (2.2.1) to Eq. (2.3.1), we observe that forces in classical mechanics are functions of only the distance between interacting bodies. We shall prove this later. Thus, the right-hand side of (2.3.1), say, the component $F_x(\Delta x)$ of the force (which is a function of Δx), is transformed into $F'_x(\Delta x')$ in accordance with (2.2.2) and remains *invariant* in different inertial reference frames.

According to the relativity principle, the left-hand side of (2.3.1) must also be invariant. As a result of Galilean transformation, the factor $\Delta \mathbf{v}/\Delta t$ is transformed into $\Delta \mathbf{v}'/\Delta t'$ and remains invariant in accordance with (2.2.4). Hence *the mass of a body is also an invariant quantity*, i.e. has the same value in different inertial reference frames. This allows us to treat mass as a mechanical property of material particles.

The state of a material particle in mechanics is defined by specifying its coordinates and velocity. The law of motion (2.3.1) establishes a connection between its states at different instants of time. If the initial coordinates and the velocity of a particle are known, as well as the forces as functions of coordinates, the entire subsequent motion of the material particle is known. The coordinates and velocity of a material particle at any instant of time can be determined with the help of Eq. (2.3.1).

It should be noted that while the mechanical law of motion (2.3.1) itself remains unchanged in all inertial reference frames, the same is not true for the trajectory of motion of a body. The trajectory of a moving body depends on its initial state.

2.4. Motion of a Material Particle in a Gravitational Field

In order to illustrate what was stated above for the law of motion in mechanics, let us consider as an example the motion of a body in a gravitational field.

The gravitational field is produced by gravitational masses. According to the modern physical concepts and experience, it can be stated that *the gravitational mass of a body is quantitatively equal to the inertial mass*, or the mass appearing in the mechanical law of motion. Consequently, we shall not make any terminological distinctions and refer simply to the mass of bodies.

The fundamental property of the gravitational field is that *all bodies, irrespective of their mass, experience the same acceleration in a given gravitational field.* In the gravitational field of the Earth, the acceleration due to gravity at small distances from the Earth is $|\mathbf{g}| = 9.8$ m/s^2. In this case, a body of mass m experiences a force of gravity equal to $m\mathbf{g}$.

Let us consider the following problem. A motor-assisted glider is catapulted from a tall cliff, and its velocity $\mathbf{v}_0$ does not change subsequently. As the glider passes over the edge of the cliff, a load is released from it. Let us first determine the motion of this load in the reference frame fixedto the glider (Fig. 2.3). The law of motion for the load has the form $m\,\Delta\mathbf{v}/\Delta t = \mathbf{F}$, where $\mathbf{F} = m\mathbf{g}$. Thus, $\Delta\mathbf{v}/\Delta t = \mathbf{g}$.

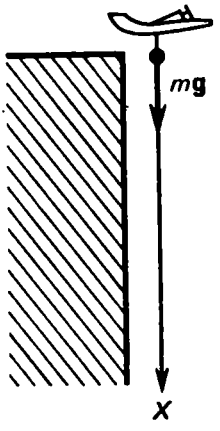


FIG. 2.3

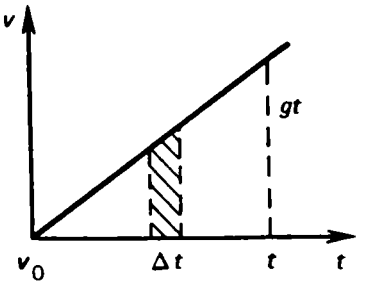


FIG. 2.4

We can write this equality in the form $\Delta \mathbf{v} = \mathbf{g}\,\Delta t$ and sum it over all small intervals of time. This gives

$$\mathbf{v} = \mathbf{g} \sum_i \Delta t_i = \mathbf{g}t.$$

This equation does not include the component $\mathbf{v}_0$ since the initial velocity of the load in the reference frame fixed to the glider is zero.

Let us find the dependence of the x-coordinate on time t. The z- and y-coordinates are not important in the present case. By definition, the velocity $v_x = \Delta x/\Delta t$. Earlier, we obtained the time dependence of the velocity in the form $v = gt$. Consequently, $\Delta x/\Delta t = gt$, or $\Delta x = gt\,\Delta t$. Graphic summation of this equality (Fig. 2.4) gives

$$x = \frac{gt^2}{2}.$$

This equation does not contain the initial coordinate which is zero.

Thus, we have obtained the time dependence of coordinates and velocity from the general form of the law of motion. Obviously, the trajectory of the load in the reference frame fixed to the glider is a vertical line since only the x-coordinate changes in this case, while $y = z = 0$.

Let us now consider the motion of the same load in the reference frame fixed to the edge of the cliff (Fig. 2.5). The vector equation of motion can be written componentwise in

the following form:

$$\left.\begin{aligned} m\frac{\Delta v_x}{\Delta t} &= F_x, \\ m\frac{\Delta v_y}{\Delta t} &= F_y. \end{aligned}\right\} \qquad \begin{aligned} &(*) \\ &(**) \end{aligned}$$

The equation (*) has already been solved. This solution and the initial conditions under which it was obtained can be written as follows:

$F_x = mg, \quad v_{x0} = 0, \quad v_x = gt,$

$x_0 = 0, \quad x = \frac{gt^2}{2}.$

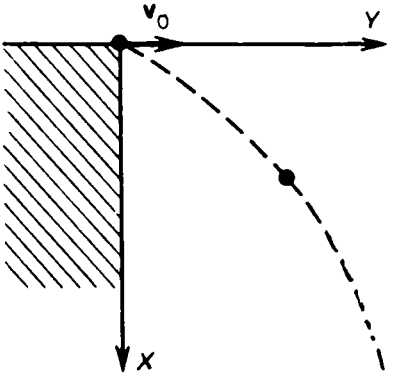


FIG. 2.5

The initial conditions for the equation of motion (**) have the form

$F_y = 0, \quad v_{y0} = v_0, \quad y_0 = 0.$

Consequently, we get

$$m\frac{\Delta v_y}{\Delta t} = 0 \rightarrow \frac{\Delta v_y}{\Delta t} = 0 \rightarrow v_y = \text{const} = v_0.$$

Relations for the y-coordinate can be written as follows:

$$\frac{\Delta y}{\Delta t} = v_0 \rightarrow \Delta y = v_0\,\Delta t \rightarrow y = v_0 t.$$

Hence we obtain

$v_y = v_0, \quad y = v_0 t.$

In this case, the equation of the trajectory is a relation between the x- and y-coordinates. In order to obtain this trajectory, let us write down the time dependence of the coordinates once again:

$$\left.\begin{aligned} x &= \frac{gt^2}{2}, \\ y &= v_0 t. \end{aligned}\right\}$$

Let us eliminate t from these equations by the method of substitution. Substituting the value of t from the second equation into the first, we obtain

$$x = \left(\frac{g}{2v_0^2}\right) y^2.$$

This is the equation of the trajectory of the load in the reference frame fixed to the edge of the cliff.

Thus, considering the motion of a load from different points of view (from the position of the pilot and of an observer standing on the cliff edge), we obtain different trajectories, viz. a vertical line and a parabola. This is due to the difference in initial conditions (velocities) for the general equations of mechanical motion.

2.5. Momentum. Law of Momentum Conservation

The law of motion

$$m \frac{\Delta \mathbf{v}}{\Delta t} = \mathbf{F}$$

considered above can be written in the following form:

$$\frac{\Delta (m\mathbf{v})}{\Delta t} = \mathbf{F} \qquad (2.5.1)$$

in view of the fact that mass is a constant quantity. The quantity $m\mathbf{v}$ is not just a formal expression but has a definite physical meaning. For a material particle, the product of its mass and velocity is called its **momentum p**. Unlike velocity, momentum obeys the law of conservation. The content of this law will be revealed below.

The student must be aware of the laboratory experiment which proves that two bodies interact with equal and opposite forces:

$$\mathbf{F}_1 = -\mathbf{F}_2,$$

or

$$\mathbf{F}_1 + \mathbf{F}_2 = 0. \qquad (2.5.2)$$

Generalizing (2.5.2) to the case of N interacting bodies, we get

$$\mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 + \ldots + \mathbf{F}_N = 0. \qquad (2.5.3)$$

This is an obvious result. Indeed, it is automatically true for a body at rest, which can be treated as an aggregate of interacting microparticles. In other words, *internal forces cannot cause the motion of a body as a whole.*

While formulating relation (2.5.3), we have actually used the concept of a **closed system**. This is a system of interacting particles on which the influence of the surroundings can be neglected. The equations of motion for N interacting material particles can be written in the form

$$\left.\begin{aligned} \frac{\Delta \mathbf{p}_1}{\Delta t} &= \mathbf{F}_1, \\ \frac{\Delta \mathbf{p}_2}{\Delta t} &= \mathbf{F}_2, \\ &\cdots\cdots \\ \frac{\Delta \mathbf{p}_N}{\Delta t} &= \mathbf{F}_N. \end{aligned}\right\} \tag{2.5.4}$$

Adding these equations and taking into account Eq. (2.5.3), we obtain

$$\frac{\Delta \sum\limits_i \mathbf{p}_i}{\Delta t} = 0.$$

This equality means that $\Delta \sum\limits_i \mathbf{p}_i = 0$, i.e.

$$\sum_{i=1}^{N} \mathbf{p}_i = \text{const}. \tag{2.5.5}$$

The expanded version of this equation has the form

$$\mathbf{p}_1 + \mathbf{p}_2 + \ldots + \mathbf{p}_N = \text{const}. \tag{2.5.6}$$

This vector equation is satisfied for a closed system of bodies. This equation means that although the momenta of individual material particles change in a closed system as a result of their interaction, *the sum of the momenta of all the particles remains constant and is independent of time* (the **law of momentum conservation**).

2.6. Law of Energy Conservation. Applications and Universal Nature of Conservation Laws

2.6.1. Law of Energy Conservation. Under the action of a force, a body moves (is displaced). In other words, work is done in moving a body. The amount A of work done is

expressed in terms of the force **F** and displacement **s** as follows:

$$A = (\mathbf{Fs}). \tag{2.6.1}$$

This relation contains the scalar product of two vectors, i.e.

$$A = |\mathbf{F}|\,|\mathbf{s}| \cos \alpha, \tag{2.6.2}$$

where α is the angle between the vectors **F** and **s**. A peculiar feature of this relation is that the work $A = 0$ for $\alpha = \pi/2$.

The capacity of systems to do work is expressed in terms of **energy**. If E_1 and E_2 represent the initial and final values of the energy of a system, the work done by the system is given by

$$A = E_2 - E_1. \tag{2.6.3}$$

This work may be positive or negative. In the former case, work is done on the system of external forces, while in the latter case, the system itself does work. Such a definition of the sign of work is in complete accord with (2.6.2).

If a body of mass m is free and moves with a velocity **v**, it is said to possess a **kinetic energy**. The expression for the kinetic energy can be obtained from dimensional analysis (see Appendix II). Indeed, the dimensions of work are $[A] = [F][s] = \mathrm{N \cdot m} = \mathrm{kg \cdot (m^2/s^2)}$, and the only possible combination of m and v having these dimensions is mv^2. The complete expression for the kinetic energy T has the form

$$T = \frac{mv^2}{2}. \tag{2.6.4}$$

By way of an illustration, let us consider the motion of a body under friction, although this problem, strictly speaking, does not lie in the framework of mechanics. This is so because mechanical energy is transformed into thermal energy during such a motion, which is not treated by mechanics. Thus, suppose that a horizontally moving body slides over a surface with an initial kinetic energy $T_1 = mv_1^2/2$, coming to rest in its final state, i.e. $v_2 = 0$. Then, according to (2.6.3), the work done against friction is given by $A = -mv_1^2/2$. If we use Eq. (2.6.2), we find that $\alpha = \pi$ since **F** and **s** are directed against each other and the work is negative.

When bodies move in an external field, the field does work on them. Let us consider, for example, bodies in a gravitational field. Suppose that a body is located at a small height x_1 over the surface of the Earth (Fig. 2.6). Naturally, work mgx_1 was done in lifting the body to this level. When left to itself, the body moves in the field of force of gravity mg and arrives at a new level x_2. The work done by the field is given by

$$A = mg(x_1 - x_2)$$

or

$$A = -(mgx_2 - mgx_1). \tag{2.6.5}$$

A comparison of (2.6.5) and (2.6.3) shows that the body must have definite values of energy at different points in the gravitational field. The energy of interacting bodies (in the present case, of the body under consideration and the Earth) is called the **potential energy**. Equation (2.6.5) can be written in the following final form:

$$A = -(U_2 - U_1), \tag{2.6.6}$$

where U is the potential energy as a function of coordinates.

It should be remarked that formula (2.6.6) is applicable not only for gravitational fields but has a more general nature. Naturally, the potential energy has different forms for different types of interaction.

The **total energy** E of a system is the sum of its kinetic energy T and potential energy U:

$$E = T + U. \tag{2.6.7}$$

In Eq. (2.6.3), we have used the total energy of the system. For closed systems, $A = 0$ and hence $E_2 = E_1$, which means that *the total energy of closed systems is conserved.* This is the essence of the **law of energy conservation.**

However, energy is conserved not only for isolated systems but also for systems located in external fields that do not vary in time. Let us consider the case of gravitational fields once again. Suppose that a body at rest ($v_0 = 0$) is initially situated at a level x_0 above the surface of the Earth (Fig. 2.7). In this state, the total energy of the body is equal to its potential energy, i.e. $E_0 = mgx_0$. Let us determine the total energy of the body at any instant of time t during its

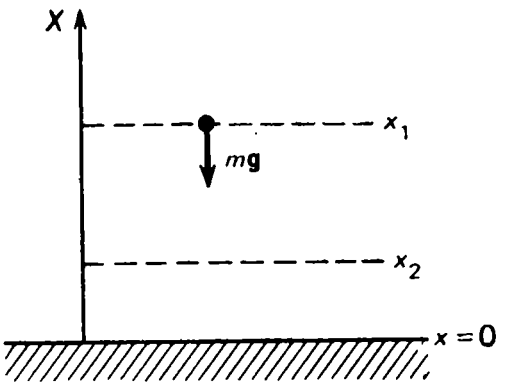


FIG. 2.6

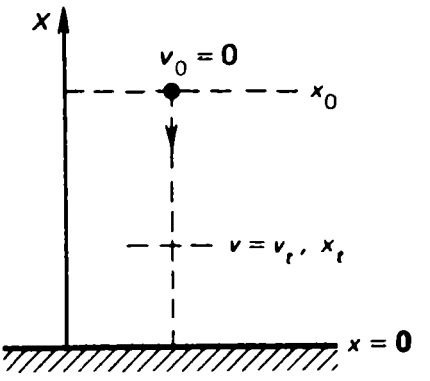


FIG. 2.7

fall. For this purpose, we need to know the coordinate x_t of the body and its velocity v_t at this instant. From formulas of Sec. 2.4, we obtain

$$\left.\begin{aligned} v_t &= gt, \\ x_t &= x_0 - \frac{gt^2}{2}. \end{aligned}\right\}$$

Substituting these values of v_t and x_t into the expression for the total energy $E = T(v) + U(x)$, we get

$$\frac{mv_t^2}{2} + mgx_t = \frac{m(gt)^2}{2} + mg\left(x_0 - \frac{gt^2}{2}\right) = mgx_0.$$

Thus, the total energy at any instant of time is equal to the energy at the initial instant of time.

2.6.2. Applications of Conservation Laws. By applying the general relations to the specific cases, we can get a better understanding of the physical content of general laws. Let us apply the conservation laws to the problems of elastic collisions of bodies and of small oscillations of bodies. These problems are quite important in physics.

Elastic collisions. Let us represent two particles by spheres of quite small size. If the collision of the particles is not accompanied by a change in their internal state, such a collision is called **elastic.** When the law of energy conservation is applied to the elastic collision of spheres, their internal energy need not be taken into consideration. Suppose that the collision of the spheres is central, i.e. takes place along the line joining their centres (Fig. 2.8). Let m_1 and

m_2 be the masses of the spheres, one of which is at rest ($v_2 = 0$), while the other moves with a velocity v_1. We have to determine the velocities u_1 and u_2 of the spheres after the collision.

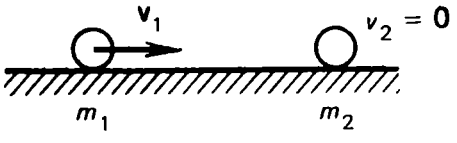


FIG. 2.8

This is a problem in two unknowns, hence we must have two equations to solve. These equations are the laws of momentum and energy conservation. We can write them in the following form:

$$\left.\begin{aligned} m_1\mathbf{v}_1 &= m_1\mathbf{u}_1 + m_2\mathbf{u}_2, \\ \frac{m_1v_1^2}{2} &= \frac{m_1u_1^2}{2} + \frac{m_2u_2^2}{2}. \end{aligned}\right\} \tag{2.6.8}$$

Since the collision of spheres is a one-dimensional process, the above system of equations assumes the form

$$\left.\begin{aligned} m_1v_1 &= m_1u_1 + m_2u_2, \\ m_1v_1^2 &= m_1u_1^2 + m_2u_2^2. \end{aligned}\right\} \tag{2.6.9}$$

In order to solve this system of equations, we transform it as follows:

$$\left.\begin{aligned} m_1(v_1 - u_1) &= m_2u_2, \\ m_1(v_1^2 - u_1^2) &= m_2u_2^2. \end{aligned}\right\} \tag{2.6.10}$$

Dividing the second of these equations by the first, we obtain

$$v_1 + u_1 = u_2. \tag{2.6.11}$$

We solve this equation simultaneously with the first of Eqs. (2.6.10). In other words, we consider the system

$$\left.\begin{aligned} v_1 + u_1 &= u_2, \\ m_1v_1 - m_1u_1 &= m_2u_2. \end{aligned}\right\} \tag{2.6.12}$$

Substituting the first of these equations into the second, we obtain

$$u_1 = \frac{m_1 - m_2}{m_1 + m_2}v_1. \tag{2.6.13}$$

Using this equation and the first of Eqs. (2.6.12), we arrive at the relation

$$u_2 = \frac{2m_1}{m_1+m_2} v_1. \qquad (2.6.14)$$

Thus, the problem formulated above is solved.

For a graphic illustration of the obtained solutions (2.6.13) and (2.6.14), let us consider one of their corollaries. Suppose that the spheres have the same mass $m_1 = m_2$. It then follows from (2.6.13) and (2.6.14) that $u_1 = 0$ and $u_2 = v_1$. In other words, a relay transmission of velocity between the spheres takes place.

In the problem considered above, the forces of interaction do not appear in explicit form. This means that the above solution of the elastic collision problem is applicable to any two objects interacting in any arbitrary manner.

By way of an example, let us consider the application of this solution to the problem of moderation of neutrons in a nuclear reactor. Matter, which does not absorb neutrons, effectively retards them only if the mass of its nuclei is close to the neutron mass. This recommendation follows from the solution obtained above. In actual practice, heavy water is one of the moderators employed in nuclear reactors. The mass of the deuterium nuclei is only twice the mass of neutrons.

Let us consider the off-centre collisions of spheres. Obviously, we must use the system (2.6.8) of vector equations in this case. For the sake of simplicity, we shall consider spheres of the same mass (Fig. 2.9). In this case, (2.6.8) assumes the form

$$\left.\begin{aligned} m\mathbf{v} &= m\mathbf{u}_1 + m\mathbf{u}_2, \\ mv^2 &= mu_1^2 + mu_2^2, \end{aligned}\right\}$$

or

$$\left.\begin{aligned} \mathbf{v} &= \mathbf{u}_1 + \mathbf{u}_2, \\ v^2 &= u_1^2 + u_2^2. \end{aligned}\right\} \qquad (2.6.15)$$

These equations are compatible only if $(\mathbf{u}_1\mathbf{u}_2) = 0$. Indeed, if we square the first equation, we obtain the second equation if this condition is satisfied. But

$$(\mathbf{u}_1\mathbf{u}_2) = u_1 u_2 \cos\alpha,$$

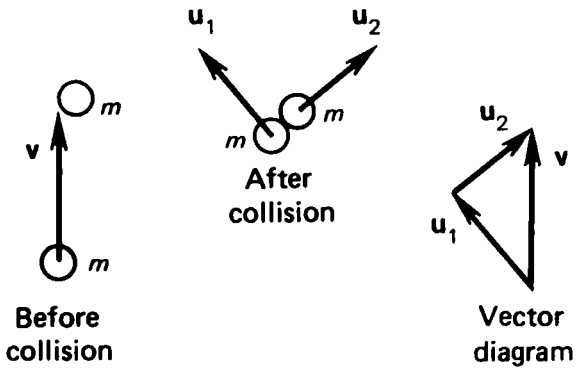


FIG. 2.9

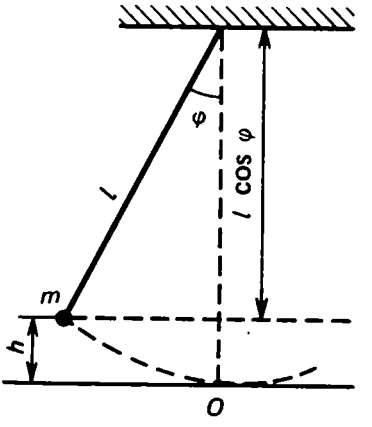


FIG. 2.10

and since $u_1 \neq u_2 \neq 0$, $\cos \alpha = 0$, i.e. $\alpha = \pi/2$. Thus, the velocity vectors for the spheres after the collision are at right angles to each other.

The system (2.6.15) of equations can be interpreted purely from a geometrical point of view (see Fig. 2.9). The vector triangle formed by the vectors $\mathbf{v}$, $\mathbf{u}_1$, and $\mathbf{u}_2$ corresponds to the first of these equations in accordance with the vector summation rule, as well as to the second equation in accordance with the Pythagorean theorem if we are dealing with a right-angled triangle.

Small oscillations. A body suspended on a thread and subjected to a single disturbance oscillates about the position of its stable equilibrium. For small angles of deviation of the pendulum from the vertical, the pendulum performs periodic oscillations for a long time when left to itself. This fact is nothing but a consequence of the law of energy conservation, as the energy of systems is conserved in a gravitational field.

Let us consider the oscillations of a material particle of mass m suspended on a thread of length l (Fig. 2.10). We write down separately the expressions for the kinetic and potential energy of the particle. The latter has the form $U(h) = mgh$, where $h = l - l \cos \varphi$, or

$$U(\varphi) = mgl(1 - \cos \varphi). \qquad (2.6.16)$$

Moreover, $U(0) = 0$, i.e. we assume that the potential energy is zero in the equilibrium position.

In the expression $T = mv^2/2$ for the kinetic energy, we introduce the **angular velocity** $\Delta\varphi/\Delta t$. Since $v = l(\Delta\varphi/\Delta t)$,

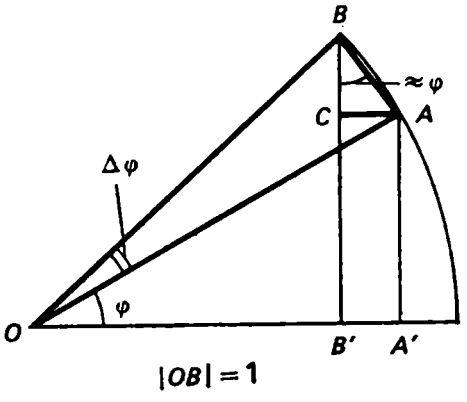


FIG. 2.11

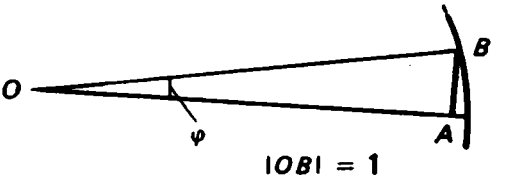


FIG. 2.12

we obtain

$$T = \frac{1}{2} ml^2 \left(\frac{\Delta\varphi}{\Delta t}\right)^2. \qquad (2.6.17)$$

The total energy of the material particle can be written as follows:

$$E = T + U = \frac{1}{2} ml^2 \left(\frac{\Delta\varphi}{\Delta t}\right)^2 + mgl\,(1 - \cos\varphi). \qquad (2.6.18)$$

The simplest periodic function that could describe the oscillations, i.e. the dependence of angle φ on time t, is the sine function

$$\varphi = \varphi_0 \sin \omega t, \qquad (2.6.19)$$

where φ_0 is the maximum angle of deviation from the equilibrium position (amplitude of the angle), and ω is called the **oscillation frequency**.

In order to find ω, we must consider some purely mathematical problems first. Let us first find the expression for $\Delta(\sin\varphi)$. By definition, $\Delta(\sin\varphi) = \sin(\varphi + \Delta\varphi) - \sin\varphi$. From the geometrical construction (Fig. 2.11) it follows that $\Delta(\sin\varphi) = |BC| = |BB'| - |AA'|$. At the same time, $|BC| = |AB| \cos\varphi$, while $|AB| \approx \Delta\varphi$. As a result, we get

$$\Delta(\sin\varphi) = \cos\varphi\,\Delta\varphi. \qquad (2.6.20)$$

We shall show that for small angles,

$$\cos\varphi \approx 1 - \frac{\varphi^2}{2}. \qquad (2.6.21)$$

Indeed, it can be seen from Fig. 2.12 that $\cos \varphi = |OA|$, but $|OA| = \sqrt{|OB|^2 - |BA|^2} \approx \sqrt{1 - \varphi^2}$. In accordance with Newton's binomial theorem, we have $(1 - \varphi^2)^{1/2} \approx 1 - \varphi^2/2$.

Let us now turn to the oscillator motion. We substitute expression (2.6.19) into (2.6.18). For this purpose, we first find the quantity $\Delta\varphi/\Delta t$. Applying the operator Δ to (2.6.19), we get $\Delta\varphi = \varphi_0 \Delta (\sin \omega t)$, with the help of Eq. (2.6.20), we write $\Delta\varphi = \varphi_0 \cos \omega t \cdot \Delta (\omega t) = \varphi_0 \omega \cos \omega t \cdot \Delta t$, or

$$\frac{\Delta\varphi}{\Delta t} = \varphi_0 \omega \cos \omega t. \tag{2.6.22}$$

Consequently, the term corresponding to the kinetic energy in Eq. (2.6.18) assumes the form

$$\frac{1}{2} ml^2 \omega^2 \varphi_0^2 \cos^2 \omega t.$$

Using (2.6.21) and (2.6.19), we obtain the following expression for the potential energy in (2.6.18):

$$\frac{1}{2} mgl\varphi_0^2 \sin^2 \omega t.$$

Combining these two expressions, we can write the total energy in the form

$$E = \frac{1}{2} ml^2 \omega^2 \varphi_0^2 \cos^2 \omega t + \frac{1}{2} mgl\varphi_0^2 \sin^2 \omega t. \tag{2.6.23}$$

Since the total energy is conserved in the system under consideration, it must be independent of time. For this condition to be satisfied, we must put

$$\omega^2 = \frac{g}{l}. \tag{2.6.24}$$

Indeed, substituting (2.6.24) into (2.6.23), we obtain

$$E = \frac{1}{2} mgl\varphi_0^2 (\cos^2 \omega t + \sin^2 \omega t),$$

or, in the final form,

$$E = \frac{1}{2} mgl\varphi_0^2. \tag{2.6.25}$$

Thus, the unknown parameter introduced in (2.6.19), viz. the oscillation frequency, is defined by (2.6.24).

2.6.3. Universal Nature of Conservation Laws. Angular Momentum. In Sec. 1.3.3 we established a relation between the symmetry properties of space and time (i.e. their homogeneity) and the inertial motion of free bodies. However, the laws of energy and momentum conservation are obeyed for free bodies, and the constancy of v in the case of inertia leads to the constancy of $m\mathbf{v}$ and $(m/2)v^2$. From this, we can naturally draw conclusions about the origin of the conservation laws: these laws must be related to the homogeneity of space and time. This circumstance explains the astonishing fact that *the laws of energy and momentum conservation are of universal nature*, i.e. they are satisfied for physical objects of quite diverse types.

In addition to homogeneity, space also has the property of **isotropy**, i.e. the physical indistinguishability of all directions in space with respect to free bodies or, to put it in a more general form, to closed systems. This symmetry of space leads to the **law of angular momentum conservation.** The angular momentum **L**, say, for a material particle, is defined as the vector product of its radius vector **r** by its momentum **p**, i.e.

$$\mathbf{L} = [\mathbf{rp}]. \tag{2.6.26}$$

Here,

$$|\mathbf{L}| = |\mathbf{r}||\mathbf{p}| \sin \alpha, \tag{2.6.27}$$

where α is the angle between the vetors **r** and **p**. The direction of the vector **L** forms a **right triplet** with the directions of the vectors **r** and **p** (Fig. 2.13). This means that when two vectors **r** and **p** are translated to a common origin, the direction of the required vector **L** will be determined by the direction of motion of a screw turned through the smaller angle between **r** and **p** (when seen from a common origin).

Let us consider some examples to illustrate the law of angular momentum conservation and its corollaries. Consider the motion of a planet in an orbit around the Sun (Fig. 2.14). In this problem, we are considering a body in an external centrally symmetric field in which the potential energy depends only on the distance from a certain point (centre). In such a field, all directions in space passing through the centre are equivalent, i.e. the angular momentum about this centre is conserved. Further, since angular momentum is a

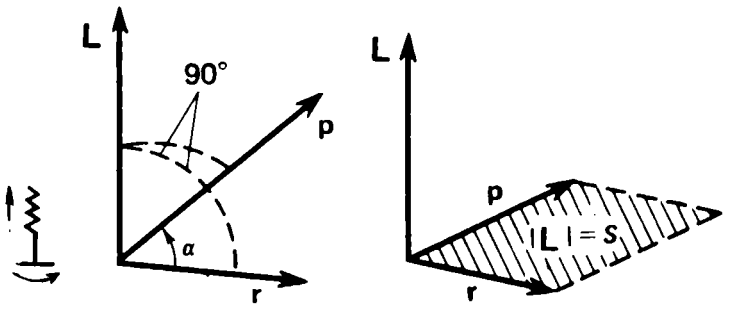


FIG. 2.13

vector quantity, its conservation means that the motion of the body must take place in a plane perpendicular to the direction of the angular momentum. In this problem, the angle between **r** and **p** is equal to $\pi/2$, i.e. the magnitude of the angular momentum can be written in the form

$$L = rp = rmv. \tag{2.6.28}$$

It follows hence that the velocity v of the motion of the body in an elliptic orbit must vary with r (see Fig. 2.14).

Using the relation

$$v = \omega r \tag{2.6.29}$$

between the linear velocity v and the angular velocity $\omega = \Delta\varphi/\Delta t$ in (2.6.28), we can write the expression for the angular momentum in the form

$$L = m\omega r^2. \tag{2.6.30}$$

Among other things, this formula explains phenomena like the increase in the angular velocity upon a contraction of rotating systems.

Before concluding this section, we observe that the practical significance of the above-mentioned conservation laws in mechanics lies in their ability to establish a relation be-

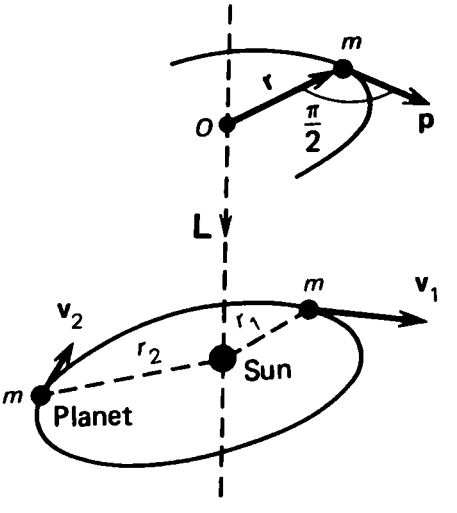


FIG. 2.14

tween quantities that are independent of the nature of interacting objects. In other words, these laws are of quite general nature.

2.7. Ultimate Velocity. Mechanics of High-Energy Particles

2.7.1. Experiments on Accelerators and Ultimate Velocity. Since the discovery of the basic laws of motion by Newton, the range of their applicability has been widened considerably. Classical mechanics superbly describes all types of motion of ordinary objects, and the classical concepts of motion are extended to objects in space. In this case also, i.e. in the planetary motion of the solar system, outstanding results are achieved. Let us turn our attention to an entirely different world, viz. thermal phenomena in which bodies (we are speaking of gases only) can be treated as aggregates of an extremely large number of molecules. Here also, a large number of concepts can be clarified by assuming that the motion of individual molecules obeys the laws of classical mechanics.

However, the triumphant march of Newtonian mechanics soon came to an end. For some phenomena, classical mechanics was found to be too coarse or simply inapplicable. In order to obtain a clear idea of the range of applicability of Newtonian mechanics, let us consider the experimental results on the acceleration of elementary particles.

The reader must be aware of the fact that at the present time huge engineering installations, viz. particle accelerators, have been constructed or are under construction in a number of countries. In these accelerators, charged particles (electrons or protons) are accelerated with the help of electromagnetic field to very high velocities. By making high-energy particles collide, the scientists try to "feel" very small spatial regions in which processes of creation and annihilation of elementary particles take place. This is a world that is enigmatic in many ways and has some profound secret hidden in it.

As the kinetic energy T of particles being accelerated gradually increases, their velocity v increases at first in accordance with the familiar relation $T = (m/2)v^2$ (Fig. 2.15). Upon a further increase in the energy of a particle, a

sharp decrease in the growth rate of its velocity is observed. At the enormous energies that have been attained in modern accelerators, the velocity of the particles practically stops increasing as it approaches the values 3×10^8 m/s. This is a sort of ultimate velocity, which we shall denote by c.

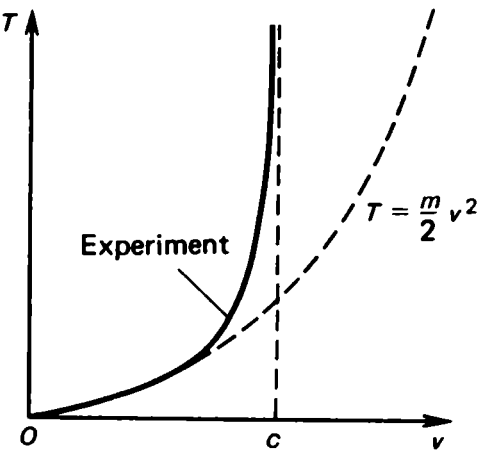


FIG. 2.15

2.7.2. Lorentz Transformations. The existence of the ultimate velocity c sharply contradicts the velocity summation rule (2.2.3) of classical mechanics. But then (2.2.3) follows from the Galilean transformations which reflect the properties of transformations of space and time upon a transition from one inertial reference frame to another. Thus, the discovery of the ultimate velocity c reveals new aspects of the properties of space and time during the motion of particles and bodies with velocities comparable with c.

Experience shows that the relativity principle remains unshakable. This means that the transformation properties (i.e. those associated with the transformations of coordinates and time) of space and time must be modified in the range of high energies. In other words, the Galilean transformations must be replaced by some other type of transformations which, however, must contain these transformations as a limiting case when the velocity $v \ll c$ (see Sec. 2.2). Of course, these new transformations must take into consideration the existence of the ultimate velocity c in nature. Such transformations have the form (see also Fig. 2.2)

$$x = \frac{x' + Vt'}{\sqrt{1 - V^2/c^2}}, \quad y = y', \quad z = z', \quad t = \frac{t' + (V/c^2)\,x'}{\sqrt{1 - V^2/c^2}}. \tag{2.7.1}$$

These are the **Lorentz transformations**, which are extremely significant in modern physics.

It follows from Eqs. (2.7.1) that (i) inertial reference frames cannot be associated with objects having a velocity

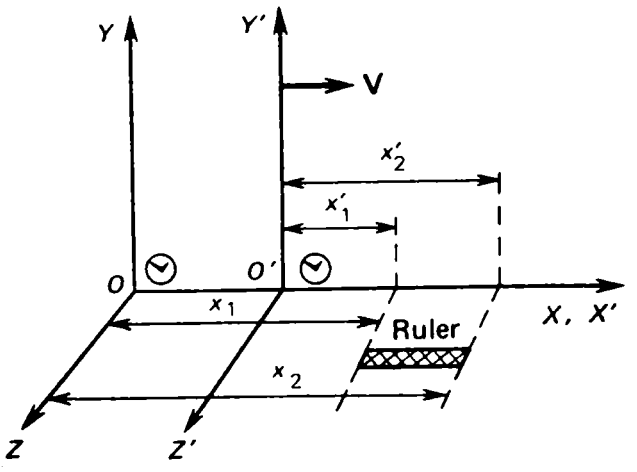


FIG. 2.16

$V = c$ ($v \equiv V$); (ii) for $V > c$, the transformations become imaginary; (iii) if $V \ll c$, (2.7.1) are transformed into (2.2.1), i.e. the laws of Newtonian classical mechanics are valid at velocities much lower than the ultimate velocity c.

The mechanics taking into consideration the ultimate velocity c is called **relativistic mechanics.**

Distances and time intervals in the theory of relativity. Let a ruler be at rest parallel to the X-axis in the inertial reference frame (X, Y, Z) (Fig. 2.16). The length of the ruler measured in this reference frame is $\Delta x = x_2 - x_1$. Let us find its length in a moving inertial reference frame (X', Y', Z'). For this purpose, we must find the coordinates of both ends of the ruler (x'_2 and x'_1) in this reference frame at the same instant of time t'. From (2.7.1), we obtain

$$x_1 = \frac{x'_1 + Vt'}{\sqrt{1 - V^2/c^2}}, \quad x_2 = \frac{x'_2 + Vt'}{\sqrt{1 - V^2/c^2}},$$

whence

$$x_2 - x_1 = \frac{x'_2 - x'_1}{\sqrt{1 - V^2/c^2}}$$

or

$$\Delta x' = \Delta x \sqrt{1 - V^2/c^2}. \tag{2.7.2}$$

From (2.7.2),! it follows that $\Delta x' \to 0$ as $V \to c$. *From the point of view of moving reference frames, the size of bodies or the distance between two points is reduced in space.* For $V \ll c$, (2.7.2) becomes the classical relation (2.2.2) $\Delta x' = \Delta x$.

Let us now suppose that the clock in the reference frame (X', Y', Z') is at rest. We consider two events occurring at the same place x' in this reference frame. The time between these two events is equal to $\Delta t' = t'_2 - t'_1$. Let us find the time Δt that has elapsed between the same events in the reference frame (X, Y, Z). From (2.7.1), we get

$$t_1 = \frac{t'_1 + (V/c^2)\,x'}{\sqrt{1 - V^2/c^2}}, \quad t_2 = \frac{t'_2 + (V/c^2)\,x'}{\sqrt{1 - V^2/c^2}},$$

whence

$$t_2 - t_1 = \Delta t = \frac{\Delta t'}{\sqrt{1 - V^2/c^2}}. \tag{2.7.3}$$

This means that *the passage of time is slowed down for moving clocks*. For $V \ll c$, (2.7.3) assumes the form (2.2.2), i.e. $\Delta t = \Delta t'$.

By way of an illustration, let us consider the experiments involving elementary particles. Most of these particles are unstable. For example, the mean lifetime of a charged π^+-meson is $\Delta t' \simeq 2.5 \times 10^{-8}$ s. This is the **intrinsic mean lifetime** of the particle, which is defined for a stationary particle. If, however, we consider a beam of π^+-mesons obtained in an accelerator, the mean lifetime Δt measured experimentally for these particles will begin to increase in accordance with (2.7.3).

Let us consider the results of specific measurements. If the mean lifetime of particles had a relativistic dependence on the state of motion, π^+-mesons would traverse a distance

$$L = (2.5 \times 10^{-8}\ \text{s}) \times (3 \times 10^8\ \text{m/s}) \simeq 7.5\ \text{m}$$

before decaying according to the scheme $\pi^+ \to \mu^+ + \nu_\mu$ (the velocity of π^+-mesons is close to the ultimate velocity c).

The following results have been obtained in experiments (Table 2.1):

Table 2.1

V	Δt, s	L, m
0	2.5×10^{-8}	
$0.9c$	5.7×10^{-8}	15.4
$0.99995c$	2.5×10^{-6}	750

Let us verify the value of Δt, say, for $V/c = 0.9$. According to (2.7.3),

$$\Delta t = \frac{\Delta t'}{\sqrt{1-V^2/c^2}}, \quad \Delta t = \frac{2.5\times 10^{-8}}{\sqrt{1-0.9^2}}\ \text{s} \simeq 5.7\times 10^{-8}\ \text{s}.$$

Table 2.1 shows that the mean lifetime of particles in a beam may increase by a factor of 10^2.

Law of velocity transformation in relativistic mechanics. Let $v_x = \Delta x/\Delta t$ be the velocity of a particle in the reference frame (X, Y, Z), and $v'_x = \Delta x'/\Delta t'$ be the velocity of the same particle in the reference frame (X', Y', Z'). We shall establish a relation between these velocities. From (2.7.1), we have

$$\Delta x = \frac{\Delta x' + V\,\Delta t'}{\sqrt{1-V^2/c^2}}, \quad \Delta t = \frac{\Delta t' + (V/c^2)\,\Delta x'}{\sqrt{1-V^2/c^2}},$$

or

$$\frac{\Delta x}{\Delta t} = \frac{\Delta x' + V\,\Delta t'}{\Delta t' + (V/c^2)\,\Delta x'}.$$

Dividing the numerator and denominator of the right-hand side of this equality by $\Delta t'$, we obtain

$$v_x = \frac{v'_x + V}{1 + v'_x V/c^2}. \qquad (2.7.4)$$

This is the **velocity summation rule in relativistic mechanics.** For $V \ll c$, (2.7.4) assumes the form $v_x = v'_x + V$ (Eq. (2.2.3)).

By way of illustration, let us consider the following examples: (i) for $v'_x = c$, we find from (2.7.4) that $v_x = c$; (ii) for $v'_x = c$ and $V = c$, it follows from (2.7.4) that $v_x = c$.

2.7.3. Relativistic Energy and Momentum. Let us derive an expression for the relativistic energy and momentum of a particle. The values **p** and E of the momentum and energy of the particle depend on the choice of the reference frame. However, there may exist a combination of **p**, E, and the relativistic constant c which does not change during a transition from one inertial reference frame to another. Such a combination of these quantities must be a scalar (number) since any scalar is invariant (constant quantity) with respect to the Lorentz transformations.

The last two of the three quantities p, E, and c are scalars, while the square of the momentum vector (p^2) is also a scalar quantity. The relation between p^2, E, and c can easily be established with the help of dimensional analysis, and has the form

$$E^2 \sim p^2c^2.$$

Consequently, we obtain the invariant

$$E^2 - p^2c^2 = (E')^2 - (p')^2c^2 = \text{const}. \tag{2.7.5}$$

Let us try to find this constant. In mechanics, the mass m of a particle is a property independent of the choice of a reference frame. Since we are considering the relativistic case of motion, the constant must also include the quantity c. Moreover, in accordance with (2.7.5), this constant must have the dimensions of the square of energy. As a result, we obtain

$$\text{const} = (mc^2)^2.$$

Thus, for a relativistic particle we arrive at a very important relation

$$E^2 - p^2c^2 = m^2c^4. \tag{2.7.6}$$

Let us analyze this equation. For $p = 0$, i.e. for a particle at rest, we obtain

$$E = mc^2. \tag{2.7.7}$$

This means that the energy of a free particle for $v = 0$ is nonzero. This is an unusual result since it is not encountered in Newtonian mechanics.

Thus, for $v \ll c$, the kinetic energy of the particle must have the form

$$E \approx mc^2 + \frac{mv^2}{2}. \tag{2.7.8}$$

Formulas (2.7.7) and (2.7.8) can be used to reconstruct the expression for the relativistic energy of the particle, viz.

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}}. \tag{2.7.9}$$

Indeed, for $v = 0$, Eq. (2.7.7) is obtained from (2.7.9). For $v \ll c$, we can apply Newton's binomial theorem to the square root in (2.7.9). This leads to relation (2.7.8) to within the terms of the order of v^4/c^4.

Finally, substituting (2.7.9) into the general relation (2.7.6), we obtain, say, for the component p_x of the momentum

$$p_x = \frac{mv_x}{\sqrt{1-v^2/c^2}}$$

or, in vector notation,

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-v^2/c^2}} . \tag{2.7.10}$$

In the limiting case $v \ll c$, this relation is reduced to the classical momentum relation $\mathbf{p} = m\mathbf{v}$.

It should be noted that the laws of energy and momentum conservation in the relativistic case can be obeyed only provided that relations (2.7.9) and (2.7.10) are valid. Relation (2.7.7) defines the **rest energy** of the particle.

Let us now consider a composite body formed by an aggregate of particles. Suppose that this body as a whole is at rest. In this case, its energy is equal to Mc^2, where M is the total mass of the body. The rest energy Mc^2 of the body is less than the total rest energy $\sum_\alpha m_\alpha c^2$ of free particles with mass m_α. In other words, the law of mass conservation is violated. However, the law of total energy conservation, including the rest energy of the particle, is satisfied.

We introduce the following definition of the **mass defect** ΔM:

$$\Delta M = \sum_\alpha m_\alpha - M. \tag{2.7.11}$$

The binding energy

$$\Delta E = -c^2 \Delta M \tag{2.7.12}$$

must be negative for stable bodies.

Later, we shall use these relations while considering nuclear fusion reactions.

2.7.4. Role of Relativistic Constant c in Physics. The discovery of the ultimate velocity c in nature set a limit to the range of applicability of the concepts of classical mechanics. The laws of Newtonian mechanics are exact and valid as long as the velocity $v \ll c$.

Equations of motion in classical mechanics were found to be invariant to the Galilean transformations. This mathematical "fact" was expressed in the form of the relativity principle applicable to the laws of motion in Newtonian mechanics. However, as a result of a transition from the Galilean to the Lorentz transformations in the region $v \sim c$ and the requirement that the relativity principle be satisfied in this region, the laws of motion must be modified at high velocities. This gives rise to a new **relativistic mechanics** whose equations of motion are invariant to the Lorentz transformations and which is automatically transformed into classical mechanics at low velocities. Among other things, it is found that as the velocity v of particles with mass $m \neq 0$ approaches c in relativistic mechanics, their energy increases indefinitely (see Eq. (2.7.9)).

Experience shows that *interactions between objects are not instantaneous, but occur with a finite, albeit very high, velocity*. This is true for both electromagnetic and gravitational interactions. Hence the concept of field is transformed into reality from the verbal form. The most astonishing fact, however, is that the velocity of propagation of interactions (fields) is equal to the ultimate velocity c. Since the velocity of propagation of a field characterizes the interaction itself, it must have the same value in all inertial reference frames. If this were not the case, the relativity principle would be violated and it would be possible to distinguish one inertial frame from another. Hence the quantity c is a universal constant **(relativistic constant)**.

The existence of the relativistic constant points towards the paradoxical properties of space and time in the range of velocities comparable with c: the size of bodies and intervals of time are found to be relative, i.e. dependent on the choice of the reference frame. Moreover, the existence of the relativistic constant c means that space and time are connected through the relation

$$c\,\Delta t = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}. \qquad (2.7.13)$$

This expression can easily be obtained if we consider that $c\,\Delta t$ is simply the distance between two closely spaced points.

The relativistic laws, which were formulated exactly by Einstein at the beginning of the 20 century, are used extensively in high-energy particle physics. The concept of field and its laws are also based on the ideas of relativity; this is supported by the fact that the relativistic constant coincides with the velocity of propagation of free electromagnetic fields in space (in particular, with the velocity of light).

Chapter 3
ELECTROMAGNETIC FIELD [1]

3.1. Electric Charge

Not all elementary particles are sources of electromagnetic field. Particles that create such a field are ascribed a definite value of the electric charge[2] e.

Like the mass m, the charge e of a particle is one of its most important properties. This is due to the fact that, like mass, charge is an invariant quantity, i.e. its value does not depend on the choice of the reference frame.

Electric charges can be of two types which we shall conditionally term positive and negative. The laws of physics do not change their form upon a replacement of all positive charges by negative, and vice versa. In other words, the physical laws are symmetric with respect to the operation $e \to -e$.

A closed system of particles obeys the **law of electric charge conservation:**

$$\sum_i e_i^+ - \sum_i e_i^- = \text{const.} \qquad (3.1.1)$$

Note that all charged elementary particles have the smallest value of charge:

$$e = 1.6 \times 10^{-19} \text{ C} \simeq 4.8 \times 10^{-10} \text{ CGSE unit of charge.}$$

The method of determining the unit of elementary charge will be described below.

---

[1] The CGS system of units is used in this chapter.

[2] This statement is not absolutely true. It is not easy to formulate statements concerning the electric charge. For example, the electric charge of a neutron is zero to a high degree of accuracy. However, the existence of a magnetic moment in a neutron indicates that this particle has a complex internal structure and that electric currents do flow in it. Consequently, the neutron possesses an intrinsic magnetic field.

3.2. Method of Field Investigation

3.2.1. Equation of Motion of a Charge in a Field. The properties of an electromagnetic field are determined from the influence exerted by it on charged particles moving in this field. The motion of charges in a field is such that it can be described as the result of the action of two quite dissimilar forces on a charge Q in space. One of these forces ($\mathbf{F}_e$) can accelerate or retard a moving charge and is independent of its velocity; the second force ($\mathbf{F}_m$) is directed at right angles to the velocity and is a function of the ratio of the charge velocity $\mathbf{v}$ to the velocity c of propagation of the electromagnetic field. Both these forces are proportional to the charge. Consequently, the point charge Q (the concept analogous to that of a material particle) is subjected to the forces

$$\mathbf{F}_e = Q\mathbf{E}, \quad \mathbf{F}_m = Q\left[\frac{\mathbf{v}}{c}\mathbf{H}\right], \tag{3.2.1}$$

where $\mathbf{E}$ and $\mathbf{H}$ are called the **electric** and **magnetic field strength** respectively.

Consequently, the equations of motion of a charge in an electromagnetic field can be written in the form

$$\frac{\Delta \mathbf{p}}{\Delta t} = Q\mathbf{E} + Q\left[\frac{\mathbf{v}}{c}\mathbf{H}\right], \tag{3.2.2}$$

where the right-hand side represents the **Lorentz force**. Significantly, (3.2.2) contains the velocity c of propagation of electromagnetic interaction, while $\mathbf{p}$, in general, is the relativistic momentum of the charged particle (see Eq. (2.7.10)).

3.2.2. Laws of Field Transformation. In accordance with the relativity principle, the equations of motion (3.2.2) must not change their form upon a transition from one inertial reference frame to another. For example, the expression for the Lorentz force in the form

$$\mathbf{F} = Q\mathbf{E} + Q\left[\frac{\mathbf{v}}{c}\mathbf{H}\right] \tag{3.2.3}$$

can be transformed only into

$$\mathbf{F}' = Q\mathbf{E}' + Q\left[\frac{\mathbf{v}'}{c}\mathbf{H}'\right]. \tag{3.2.4}$$

Obviously, for this purpose, the law of field transformation

$$\mathbf{E} \to \mathbf{E}', \quad \mathbf{H} \to \mathbf{H}'$$

must be quite definite. Thus, in the case $V \ll c$, where V is the relative velocity of two inertial reference frames (see Fig. 2.16), this transformation has the following form (accurate to the terms of the order of V/c):

$$\mathbf{E} = \mathbf{E}' + \frac{1}{c}[\mathbf{H}'\mathbf{V}], \quad \mathbf{H} = \mathbf{H}' - \frac{1}{c}[\mathbf{E}'\mathbf{V}], \tag{3.2.5}$$

while $\mathbf{v} = \mathbf{v}' + \mathbf{V}$.

Let us verify this. Substituting (3.2.5) into (3.2.3), we obtain

$$\mathbf{F} = Q\left(\mathbf{E}' + \frac{1}{c}[\mathbf{H}'\mathbf{V}]\right) + Q\left[\frac{\mathbf{v}'+\mathbf{V}}{c}\left(\mathbf{H}' - \frac{1}{c}[\mathbf{E}'\mathbf{V}]\right)\right]$$

$$= Q\mathbf{E}' + \frac{Q}{c}[\mathbf{H}'\mathbf{V}] + Q\left[\frac{\mathbf{v}'}{c}\,\mathbf{H}'\right] + \frac{Q}{c}[\mathbf{V}\mathbf{H}']$$

$+$ small terms of the order $(V/c)^2$.

Considering that the vector product changes sign upon a transposition of vectors, the relation obtained above can be reduced to (3.2.4).

Let us consider some properties of transformations (3.2.5). If the magnetic field strength $\mathbf{H}' = 0$ in some reference frame (X', Y', Z'), the electric and magnetic fields in the reference frame (X, Y, Z) are connected, in accordance with (3.2.5), through the relation

$$\mathbf{H} = \frac{1}{c}[\mathbf{V}\mathbf{E}]. \tag{3.2.6}$$

If, however, the electric field strength $\mathbf{E}' = 0$ in the reference frame (X', Y', Z'), we obtain the following relation for the reference frame (X, Y, Z):

$$\mathbf{E} = -\frac{1}{c}[\mathbf{V}\mathbf{H}]. \tag{3.2.7}$$

Thus, *electric and magnetic fields are relative*, i.e. have different properties in different reference frames. For example, the electric or magnetic field strength may be zero in one reference frame and nonzero in another.

The field transformations (3.2.5) generalized to the case of velocities V comparable with c assume the following form of electric and magnetic field transformation relations:

$$E_x = E'_x, \quad E_y = \frac{E'_y + (V/c) H'_z}{\sqrt{1 - V^2/c^2}}, \quad E_z = \frac{E'_z - (V/c) H'_y}{\sqrt{1 - V^2/c^2}}, \tag{3.2.8}$$

$$H_x = H'_x, \quad H_y = \frac{H'_y - (V/c) E'_z}{\sqrt{1 - V^2/c^2}}, \quad H_z = \frac{H'_z + (V/c) E'_y}{\sqrt{1 - V^2/c^2}}. \tag{3.2.9}$$

The limiting transition in these formulas to $V \ll c$ leads to Eqs. (3.2.5) written componentwise.

3.3. Laws of Electromagnetic Field

3.3.1. New Objects and New Mathematics. The state of a material particle was described by specifying its position in space and its velocity. Such an approach is untenable for electromagnetic fields. A field is defined only if the values of the vectors **E** and **H** are known at each point in space. In the general case, **E** and **H** are functions of not only coordinates **r** but also time t, i.e. **E** (**r**, t) and **H** (**r**, t).

In the description of vector fields, including electromagnetic fields, the concepts of **flux** and **circulation of the field** are found to be very convenient. If we mentally draw a certain surface S in a vector field **A** (**r**) (Fig. 3.1), the vector **A** must have a definite value and orientation at each point on this surface.

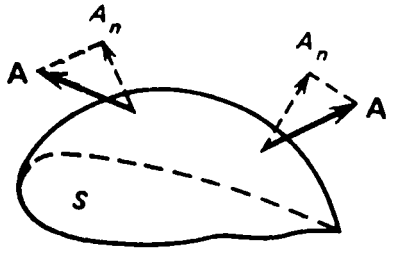


FIG. 3.1

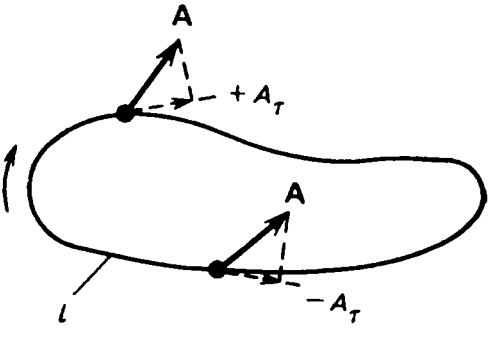


FIG. 3.2

Let us consider for each vector **A** its projection A_n on the outward normal to the surface drawn from the initial point of the vector. We can then write[3]

$$\begin{pmatrix}\text{Flux}\\ \text{of the vector field } \mathbf{A}(\mathbf{r})\\ \text{across the surface } S\end{pmatrix} = \begin{pmatrix}\text{Average}\\ \text{normal component}\\ \text{of the vector } \mathbf{A}(\mathbf{r})\end{pmatrix} \times \begin{pmatrix}\text{Area } S\\ \text{of the surface}\end{pmatrix}$$

$$= \langle A_n \rangle S. \tag{3.3.1}$$

In order to obtain the concept of circulation of a vector field, we mentally draw a certain closed contour l in the field **A** (**r**). Each point of the contour has a vector **A** corresponding to it (Fig. 3.2). For each vector **A**, we consider its projection A_τ on the tangent to the contour (taking into account its sign with respect to the direction of circumvention). We can express the circulation of the field in the following form:

$$\begin{pmatrix}\text{Circulation}\\ \text{of the vector field } \mathbf{A}(\mathbf{r})\\ \text{along the closed contour } l\end{pmatrix} = \begin{pmatrix}\text{Average}\\ \text{tangential component}\\ \text{of the vector } \mathbf{A}(\mathbf{r})\end{pmatrix} \times \begin{pmatrix}\text{Length } l\\ \text{of the path}\\ \text{of circumvention}\end{pmatrix}$$

$$= \langle A_\tau \rangle l. \tag{3.3.2}$$

We shall use the concepts of flux and circulation to formulate the basic laws of electromagnetic field and, wherever possible, endeavour to substantiate them experimentally.

3.3.2. First Field Equation. Relation Between Electric Field and Electric Charge. The first field equation can be formulated as follows (Fig. 3.3):

$$\begin{pmatrix}\text{Flux of the electric field } \mathbf{E}\\ \text{across any}\\ \text{closed surface } S\end{pmatrix} = 4\pi \begin{pmatrix}\text{Charge}\\ \text{within}\\ \text{the closed surface}\end{pmatrix},$$

[3] Here and below, angle brackets will be used to denote the mean values of quantities.

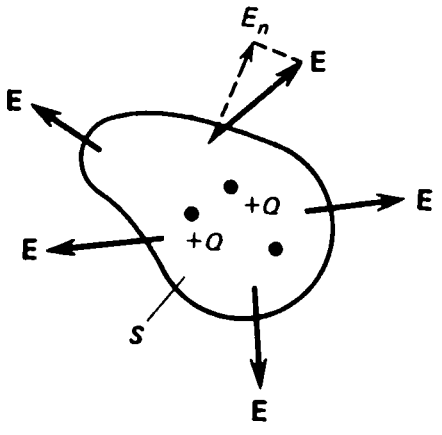


FIG. 3.3

or

$$\langle E_n \rangle S_{cl} = 4\pi \sum_i Q_i. \qquad (3.3.3)$$

This equation indicates that *charge is a source of electric field.*

For a visual interpretation of Eq. (3.3.3), we introduce the concept of **lines of force of the field** (or **field lines**) defined as the lines the tangent to which at each point coincides with the direction of the field. The density of the field lines is proportional to the field strength. In this connection, it is useful to recall the laboratory experiments with plumes[4] (Fig. 3.4). These experiments lead to the conclusion that *electric field lines start and terminate at the charges or extend to infinity.*

3.3.3. Second Field Equation. Absence of Magnetic Charges. This equation has the form (Fig. 3.5)

$$\begin{pmatrix}\text{Flux of the magnetic field } \mathbf{H} \\ \text{across any} \\ \text{closed surface } S\end{pmatrix} = 0,$$

or

$$\langle H_n \rangle S_{cl} = 0. \qquad (3.3.4)$$

This equation indicates that *there are no magnetic charges in nature. Magnetic field lines do not start or terminate at any point and are always closed.*

3.3.4. Third Field Equation. Relation Between Current and "Something" with a Vortex Magnetic Field. This equation is

[4] The experiments are carried out by using a bunch of narrow paper strips which are fixed at one end to the tip of an insulating support. The electric charge is transferred to the strip through a charged rod or from a Wimshurst (electrostatic) machine. In an isolated charged plume, paper strips are arranged radially. The mutual configuration of strips in the space between a pair of charged plumes (having like or unlike charges) gives an idea about the nature of electric field.

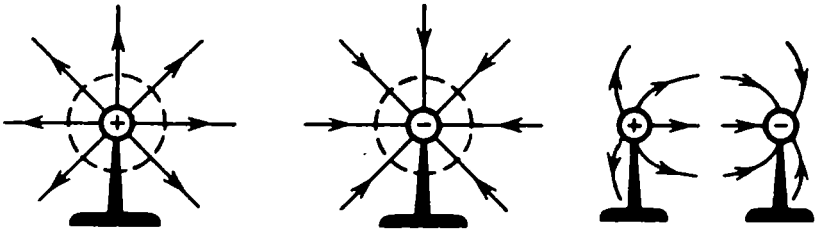


FIG. 3.4

formulated as follows (Fig. 3.6):

$$\begin{pmatrix}\text{Circulation}\\ \text{of the magnetic field } \mathbf{H}\\ \text{along an arbitrary}\\ \text{closed contour } l\end{pmatrix} = \begin{pmatrix}\text{Electric current}\\ \text{across the surface } S\\ \text{stretched}\\ \text{on the contour } l\end{pmatrix} + (\text{A term hitherto unknown}),$$

or

$$\langle H_\tau \rangle\, l = \frac{4\pi}{c}\,(\langle j_n \rangle\, S) + (\text{"something"}). \qquad (3.3.5)$$

This equation contains the normal component of the **electric current density** vector $\mathbf{j} = \rho\mathbf{v}$, where ρ is the number of charges per unit volume, and $\mathbf{v}$ is the velocity of a charge at a given point on the surface.

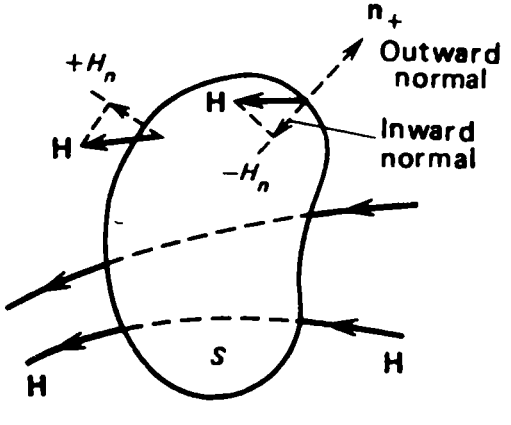


FIG. 3.5

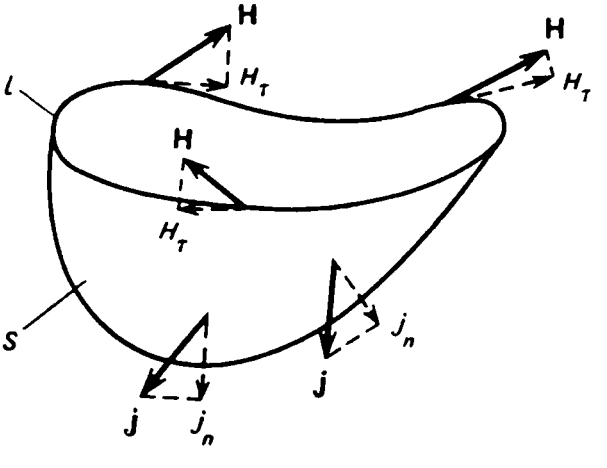


FIG. 3.6

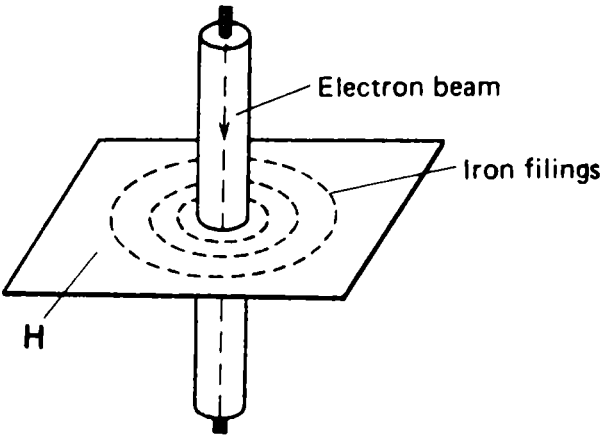


FIG. 3.7

It can be seen from Eq. (3.3.5) that *moving electric charges serve as a source of a vortex magnetic field.* But an analysis of the motion of charges requires that the relativistic constant c be taken into account. The nature of the unknown term in (3.3.5) will be revealed later.

Equation (3.3.5) can be experimentally verified by using a cathode-ray tube (Fig. 3.7). Upon the passage of an electron beam, iron filings are grouped together in concentric circles along the lines of the magnetic field **H**. The very fact of the formation of an electron beam in the tube means that the charges are accelerated in an electric field according to the law $\mathbf{F}_e = e\mathbf{E}$.

The existence of the magnetic component $\mathbf{F}_m = e[(\mathbf{v}/c)\mathbf{H}]$ of the Lorentz force can be experimentally observed on a dual-trace oscillograph (Fig. 3.8). By switching on each electron gun separately, we mark the position

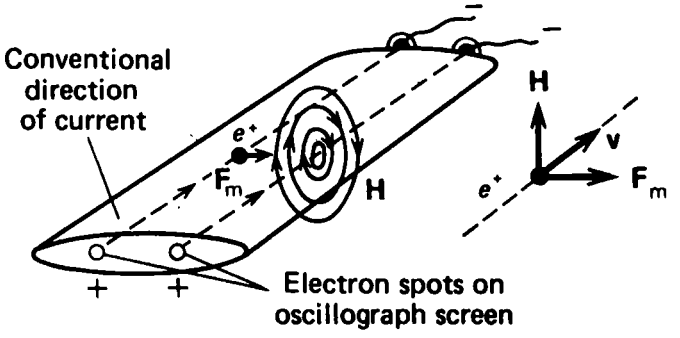


FIG. 3.8

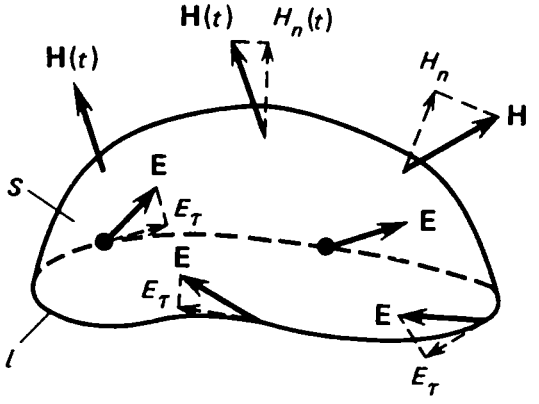


FIG. 3.9

of the luminous spots on the screen. If, however, both beams are simultaneously emitted, the spots become closer (for the sake of convenience, in using the left-hand rule in Fig. 3.8, electrons e^- are replaced by oppositely charged particles e^+ moving in the opposite direction).

3.3.5. Fourth Field Equation. Relation Between a Varying Magnetic Field and a Vortex Electric Field. This equation has the form (Fig. 3.9):

$$\begin{pmatrix}\text{Circulation}\\ \text{of the electric field } \mathbf{E}\\ \text{along the contour } l\end{pmatrix} = -\frac{1}{c}\frac{\Delta}{\Delta t}\begin{pmatrix}\text{Flux}\\ \text{of the magnetic field } \mathbf{H}\\ \text{across the surface } S\\ \text{stretched on the contour } l\end{pmatrix},$$

or

$$\langle E_\tau\rangle\, l = -\frac{1}{c}\frac{\Delta}{\Delta t}(\langle H_n\rangle\, S). \qquad (3.3.6)$$

In this equation, the magnetic field is a function of time. For varying fields, it is always necessary to consider the velocity c of propagation of the fields. It follows from this equation that a *varying magnetic field can produce a vortex electric field.*

This equation can be illustrated quite clearly in particle accelerators like betatron (Fig. 3.10). A vacuum tube is placed in a rapidly increasing magnetic field. This results in the

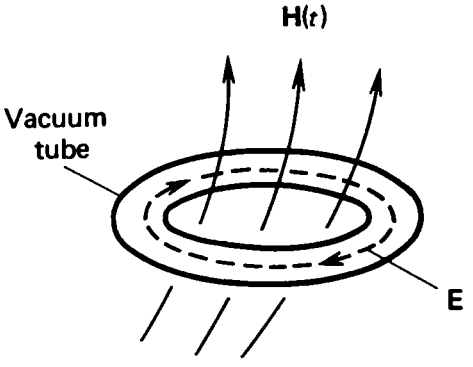


FIG. 3.10

emergence of a vortex electric field in the tube, accelerating the electrons injected into the tube.

3.3.6. Additional Analysis of the Third Field Equation. Relation Between a Varying Electric Field and a Vortex Magnetic Field. Fields **E** and **H** are functions of coordinates **r** and time t. The system of equations (3.3.3)-(3.3.6) is complete if it can be used to determine **E** (**r**, t) and **H** (**r**, t). Equations (3.3.3) and (3.3.4) give $\langle E_n \rangle$ and $\langle H_n \rangle$ as functions of coordinates, while Eqs. (3.3.5) and (3.3.6) describe the spatial distribution of $\langle H_\tau \rangle$ and $\langle E_\tau \rangle$. The time dependence of $\langle H_n \rangle$ appears in Eq. (3.3.6). Consequently, the unknown term in (3.3.5) must determine the time dependence of $\langle E_n \rangle$.

Let us determine the form of the unknown term in Eq. (3.3.5). To begin with, all the time-dependent expressions in (3.3.3)-(3.3.6) naturally contain c (the velocity of propagation of the electromagnetic field). From the considerations of symmetry and dimensional analysis, it can be concluded that this term must have the form

$$\frac{1}{c}\frac{\Delta}{\Delta t}(\langle E_n \rangle S).$$

Thus, the third field equation can be written in the following final form:

$$\langle H_\tau \rangle l = \frac{4\pi}{c}(\langle j_n \rangle S) + \frac{1}{c}\frac{\Delta}{\Delta t}(\langle E_n \rangle S). \qquad (3.3.7)$$

This equation is graphically illustrated in Fig. 3.11. It follows from Eq. (3.3.7) that a *vortex magnetic field is produced not only by current (moving charges) but also by a varying electric field.*

The term $(1/c)\Delta(\langle E_n \rangle S)/\Delta t$ in (3.3.7) is clearly manifested in experiments involving a.c. circuits containing capacitors. These experiments will be described in Chap. 7.

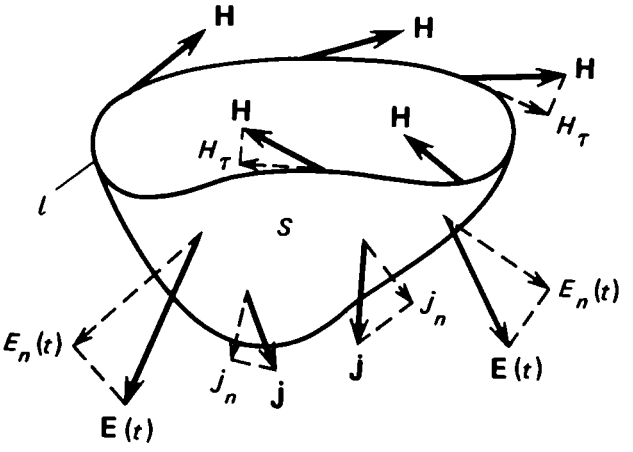


FIG. 3.11

3.3.7. Maxwell's Field Equations. The four field equations (3.3.3), (3.3.4), (3.3.6), and (3.3.7) constitute a single system called **Maxwell's (field) equations.** These equations completely define an electromagnetic field if we know the arrangement and motion of electric charges. These equations, which must be supplemented by formula (3.2.3) for the Lorentz force, express the most general laws of electromagnetic phenomena and describe the entire range of electromagnetism.

It is interesting to note that Maxwell's equations contain the law of charge conservation. Maxwell's equations and Eq. (3.2.2) describing the motion of a charge in an electromagnetic field can be used for deriving an expression for the energy and momentum of the field. *The total momentum and the total energy of the particle and the field are conserved in the closed particle-field system.*

Maxwell's equations satisfy the relativity principle: they are invariant to the Lorentz transformations (see Eqs. (2.7.1)). It should be noted that historically, the latter were obtained from the requirement of invariance of the electromagnetic field equations in various inertial reference frames.

A significant feature of electromagnetic fields is that they obey the **superposition principle**: *the field resulting from an aggregate of moving charges can be represented as the (vector) sum of the fields of the individual charges.* In other words, when fields are superimposed, their mutual influence can be ne-

glected. Maxwell's equations satisfy the superposition principle.

Our next problem is to consider a number of corollaries of the general laws of electromagnetism and their applications. These applications make the general laws more illustrative and tangible.

3.4. Constant Electric Field

A constant electric field must be described by the first field equation (3.3.3).

3.4.1. Field of a Stationary Point Charge. This field has an obvious central symmetry with the centre at the point charge Q (Fig. 3.12). The equation $E_n = E$ must be satisfied at all points on the spherical surface of area S mentally drawn in the field of the point charge. Since E is the same at all points of a sphere of radius r, we can write $\langle E_n \rangle = E$. Consequently, the field equation (3.3.3) can be presented in the form $ES = 4\pi Q$ or $E \cdot 4\pi r^2 = 4\pi Q$, i.e.

$$E = \frac{Q}{r^2} . \tag{3.4.1}$$

This is the electrostatic field of a stationary point charge. The validity of this relation has been experimentally proved for electrons right down to distances of the order of 10^{-16} m.

3.4.2. Field of Charges Distributed over a Sphere, Line or Plane Surface. Let us find the electric field **E** for a charged sphere of radius R having an overall charge Q. In view of the spherical symmetry of the field in this case, we can construct two auxiliary surfaces: the first surface S_1 has a radius r and is located outside the charged sphere, while the second surface S_2 lies inside the sphere containing the charge Q (Fig. 3.13). For the surface S_1, the field equation (3.3.3) can be written in the form $E \cdot 4\pi r^2 = 4\pi Q$, i.e.

$$E = \frac{Q}{r^2} . \tag{3.4.2}$$

For the surface S_2, we obtain from (3.3.3) $ES_2 = 0$ since there is no charge inside S_2. Thus, *there is no field inside a charged sphere.*

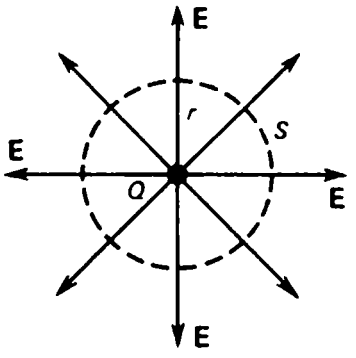


FIG. 3.12

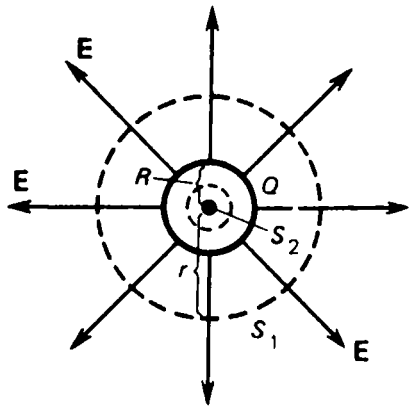


FIG. 3.13

Let us compute the value of **E** for charges lying along a straight line. Let ρ be the linear density of charge distribution. The electric field in this problem has axial symmetry, and the symmetry axis is directed along the line containing the charges (Fig. 3.14). We construct a closed cylindrical surface in which S_1 are the areas of the upper and lower bases, while S_2 is the area of the lateral surface. The flux of **E** across the surfaces S_1 is zero, while the flux across the surface S_2 is given, in accordance with Eq. (3.3.3), by $E \cdot 2\pi r \cdot h = 4\pi\rho h$, where h is the height of the cylindrical surface, and r is its radius. Consequently, we get

$$E = \frac{2\rho}{r}. \tag{3.4.3}$$

Let us now find the electric field **E** produced by charges located in a plane with the surface charge density σ

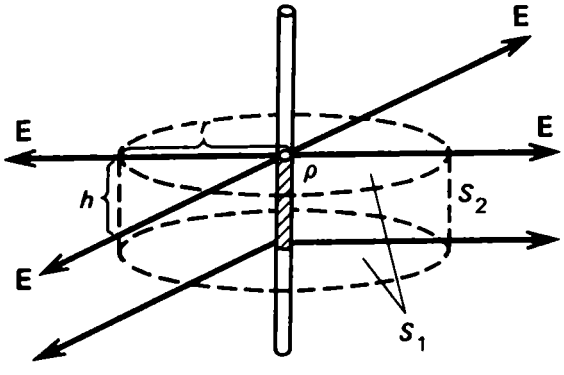


FIG. 3.14

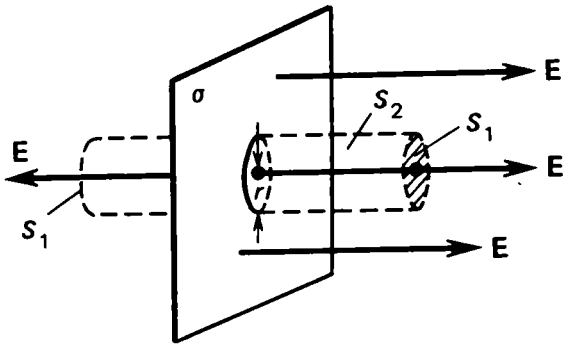


FIG. 3.15

(Fig. 3.15). It follows from symmetry considerations that the electric field vector **E** is at right angles to the plane. In order to determine **E**, we construct an auxiliary closed cylindrical surface with the bases S_1 and the lateral surface S_2. The vector **E** slides along the generator of the cylinder, and its flux across S_2 is zero. The flux across S_1 is defined, in accordance with (3.3.3), as $E \cdot 2\pi r^2 = 4\pi\sigma \cdot \pi r^2$, i.e.

$$E = 2\pi\sigma. \tag{3.4.4}$$

The field of a uniformly charged infinite plane is uniform all over the space and has opposite directions on both sides of the plane.

Finally, let us determine the field **E** produced by unlike charges situated on parallel infinite planes under the condition that the surface charge density σ is the same in both cases (Fig. 3.16). The field lines of **E** emerge from the positively charged surface and enter the negatively charged surface (in Fig. 3.16, $\mathbf{E}^+$ and $\mathbf{E}^-$ denote the field produced by the positively and negatively charged surface respectively). It can be seen from the figure that the field is completely sup-

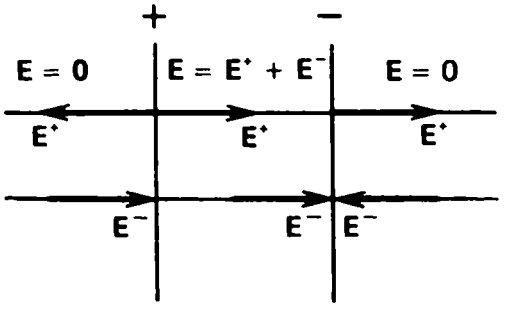


FIG. 3.16

pressed outside the surfaces and is doubled between the surfaces. Taking (3.4.4) into consideration, we obtain for the "capacitor" the relation $E = 2 \cdot 2\pi\sigma$, or

$$E = 4\pi\sigma. \tag{3.4.5}$$

3.4.3. Electrostatic Energy of Charges. Field Potential. We shall use the following method for determining the energy of interaction between two charges: a charge Q_2 is located in the electric field E produced by a stationary charge Q_1. In this case, the electric field acts on the charge Q_2 by a force $F = Q_2E$ (see Eq. (3.2.1)). Using expression (3.4.1) for E, we obtain

$$F = \frac{Q_1Q_2}{r^2}, \tag{3.4.6}$$

where r is the separation between the point charges. This relation is called **Coulomb's law** which is found to be convenient for defining a unit of charge.[5]

Knowing the force of interaction (3.4.6) between charges, we can find the work done in displacing the charge Q_2 in the field of the stationary charge Q_1. We start with the general expression for an element of work:

$$\Delta A = F\Delta r. \tag{3.4.7}$$

The total work done in displacing the charge from r_0 to r (Fig. 3.17) is then defined as

$$A = \sum F(r)\,\Delta r = Q_1Q_2 \sum \frac{\Delta r}{r^2}, \tag{3.4.8}$$

where r is the running coordinate and its final value.

Thus, the problem is reduced to determining the sum $\sum (\Delta r/r^2)$. We split the entire segment between r_0 and r into smaller segments (see Fig. 3.17) and introduce the quantity $\sqrt{r_1r_0}$, i.e. the geometric mean distance between points r_1 and r_0. As a result, we obtain

$$\sum \frac{\Delta r}{r^2} = \frac{r_1 - r_0}{r_1r_0} + \frac{r_2 - r_1}{r_2r_1} + \frac{r_3 - r_2}{r_3r_2} + \frac{r - r_3}{rr_3}$$

[5] This is the approach used in the CGS system. In SI units, the unit of charge is defined in terms of a unit of current, viz. as [A·s]. The ampere standard is introduced through the magnetic interaction of currents.

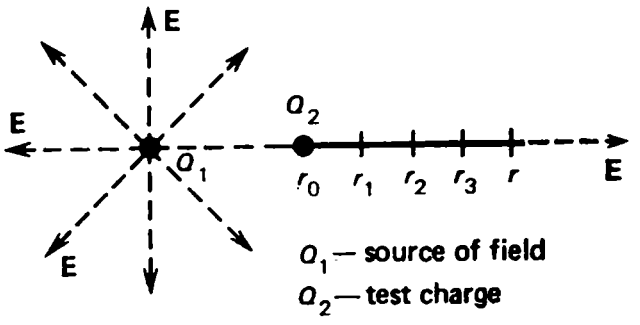


FIG. 3.17

or

$$\sum \frac{\Delta r}{r^2} = \left(\frac{1}{r_0} - \frac{1}{r_1}\right) + \left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$
$$+ \left(\frac{1}{r_2} - \frac{1}{r_3}\right) + \left(\frac{1}{r_3} - \frac{1}{r}\right) = \frac{1}{r_0} - \frac{1}{r} .$$

With the help of this result, we can present (3.4.8) in the form

$$A = Q_1 Q_2 \left(\frac{1}{r_0} - \frac{1}{r}\right) ,$$

or

$$A = -\left(\frac{Q_1 Q_2}{r} - \frac{Q_1 Q_2}{r_0}\right) . \qquad (3.4.9)$$

Comparing formula (3.4.9) with the general expression (2.6.12), we arrive at the following relation for the electrostatic energy of interaction between the charges:

$$U = \frac{Q_1 Q_2}{r} . \qquad (3.4.10)$$

Here, the charges Q_1 and Q_2 are separated by a distance r. As $r \to \infty$, the potential energy $U \to 0$.

It follows from (3.4.10) that the potential energy may have different signs: it will be positive if both charges are of the same sign and negative otherwise. This explains the fact that we observe a repulsion as well as an attraction of charges (Fig. 3.18).

We introduce the concept of **field potential**:

$$\varphi = \frac{U}{Q_2} . \qquad (3.4.11)$$

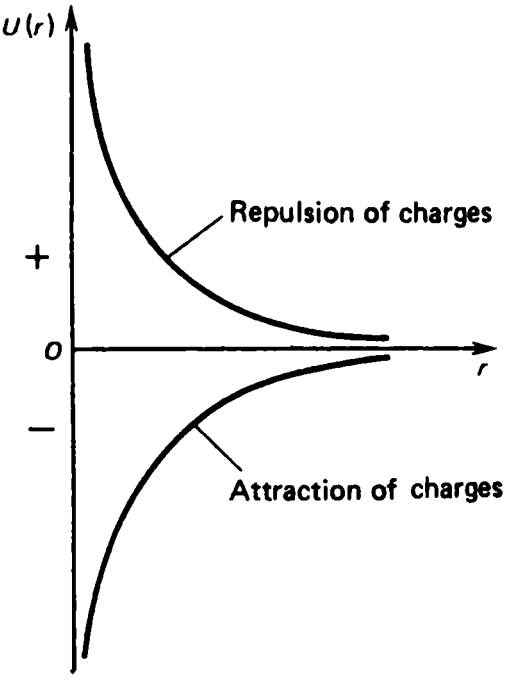


FIG. 3.18

The potential of a field characterizes its energy. It is equal to the energy possessed by a unit test charge placed at a given point in the field. Substituting (3.4.10) into (3.4.11), we obtain

$$\varphi = \frac{Q_1}{r} . \tag{3.4.12}$$

This is the expression for the potential of the field produced by the charge Q_1 at a point separated from the charge by a distance r.

The work (3.4.9) done in displacing a test charge Q_2 from point r_0 to point r in a field with potential (3.4.12) can be presented in the form

$$A = -Q_2 [\varphi (r) - \varphi (r_0)]. \tag{3.4.13}$$

The potential φ of a field can easily be connected with its strength **E**. Indeed, $A = -Q_2 \Delta\varphi$, on the one hand, and $A = Q_2 (\mathbf{E} \Delta\mathbf{r})$, on the other. If $\mathbf{E} \parallel \Delta\mathbf{r}$, we obtain

$$E = -\frac{\Delta\varphi}{\Delta r} , \tag{3.4.14}$$

where $\Delta\varphi/\Delta r$ is the potential drop over a unit length.

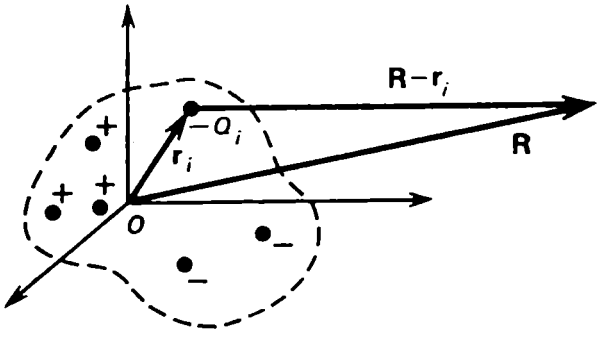


FIG. 3.19

This relation allows us to use two methods for defining an electric field, viz. in terms of potential or in terms of the field strength vector.

3.4.4. Field of a Dipole. Charge-Dipole and Dipole-Dipole Interactions. In some molecules, the atoms constituting them are arranged asymmetrically. This means that the distribution of both types of charge in a molecule is also asymmetric. Although the molecule is electrically neutral on the whole, the incoincidence of the centres of positive and negative charges results in an electric field outside the molecule. Let us calculate this field.

Suppose that we have a certain system of charges. We fix the origin O of coordinates somewhere in this system (Fig. 3.19) and calculate the potential of the field at a point far removed from the system under consideration. We denote by $\mathbf{r}_i$ the radius vector of a charge Q_i, by $\mathbf{R}$ the "coordinates" of the point of observation, and by $|\mathbf{R} - \mathbf{r}_i|$ the distance between the charge Q_i and the point where the field is being calculated. Then, in accordance with (3.4.12), we obtain the following expression for the potential of the field produced by the charge Q_i at the point under consideration:

$$\varphi_i = \frac{Q_i}{|\mathbf{R}-\mathbf{r}_i|} . \qquad (3.4.15)$$

Let us analyze this relation for $R \gg r_i$. We shall obtain an approximate expression for $(R - r_i)^{-1}$. For this purpose, we use the binomial theorem $(a - b)^n \approx a^n - na^{n-1}b$, where we must put $n = -1$. This gives

$$(R-r_i)^{-1} \approx \frac{1}{R} + \frac{r_i}{R^2} .$$

The substitution of this expression into (3.4.15) gives

$$\varphi_i = \frac{Q_i}{R} + \frac{Q_i r_i}{R^2} .$$

In order to determine the potential of the field produced at the point of observation by all charges of the system, we use the superposition principle: $\varphi = \sum \varphi_i$, i.e.

$$\varphi = \frac{\sum Q_i}{R} + \frac{\sum Q_i |\mathbf{r}_i|}{R^2} . \qquad (3.4.16)$$

Since $\sum Q_i = 0$ in the system of charges under consideration, we can write (3.4.16) as follows:

$$\varphi = \frac{|\mathbf{d}|}{R^2} , \qquad (3.4.17)$$

where $\mathbf{d} = \sum Q_i \mathbf{r}_i$ is called the **electric dipole moment** of the system of charges.

Let us determine the energy of interaction of a unit charge with a dipole system. For this purpose, we consider a charge Q in the field φ of the dipole. Using (3.4.11) and (3.4.17), we can write

$$U = \frac{Q|\mathbf{d}|}{R^2} . \qquad (3.4.18)$$

The energy of interaction of two dipole systems having zero total charge each, electric dipole moments $\mathbf{d}_1$ and $\mathbf{d}_2$, and separated by a distance larger than their size can be written in the form

$$U \approx \frac{(\mathbf{d}_1 \mathbf{d}_2)}{R^3} . \qquad (3.4.19)$$

Finally, let us write the formulas for the energies of interaction of the charge-charge, charge-dipole, and dipole-dipole systems:

$$U_{QQ} = \frac{Q^2}{R} , \quad U_{Qd} = \frac{Q|\mathbf{d}|}{R^2} , \quad U_{dd} \approx \frac{\mathbf{d}^2}{R^3} . \qquad (3.4.20)$$

These formulas will be used to analyze the interaction of molecules.

3.5. Constant Magnetic Field

Constant magnetic fields must be studied on the basis of the third field equation (3.3.7) in which the time-dependent terms are omitted, and the current is assumed to be stationary. As a result, (3.3.7) becomes

$$\langle H_\tau \rangle l = \frac{4\pi}{c} (\langle j_n \rangle S). \tag{3.5.1}$$

We shall ccnsider some symmetric systems of this type.

3.5.1. Magnetic Field of a Direct Current. Suppose that charges move in space along an infinite straight line, and that the **current density j** (i.e. the charge transferred per second across a unit area in a direction perpendicular to the area element) remains constant. For a contour l, we choose a circle in a plane perpendicular to the current I (Fig. 3.20). The surface of area S is bounded by the contour l. Considering the symmetry and the vortex nature of the magnetic field **H**, we can conclude that the vector **H** is a function of r alone and is directed along the tangent to the contour l, i.e. $\langle H_\tau \rangle = |\mathbf{H}|$. Consequently, (3.5.1) assumes the form $H \cdot 2\pi r = (4\pi/c) I$, where $I = jS$ is the current. Finally, for the magnetic field produced by a direct current, we obtain

$$H(r) = \frac{2}{c} \frac{I}{r} . \tag{3.5.2}$$

3.5.2. Magnetic Field of a Current Surface. Suppose that a current flows in an unlimited plane, the linear current density being equal to i. We draw an auxiliary contour l in a plane perpendicular to that of the current (Fig. 3.21). Symmetry considerations indicate that the magnetic field vector **H** is parallel to the current surface. Since the latter is infinite, **H** cannot depend on the distance from it. In other words, *the magnetic field on both sides of a current surface is uniform, and the directions of* **H** *on opposite sides of this surface are antiparallel.* This follows from the fact that magnetic field lines must converge at infinity.

Let us apply Eq. (3.5.1) to the present case. Circumventing the contour l in the clockwise direction relative to the direction of the current, we obtain $\langle H_\tau \rangle l = H \cdot l_1 + 0 \cdot 2l_2 + H \cdot l_1 + 0 \cdot 2l_2 = H \cdot 2l_1$. The total current across the con-

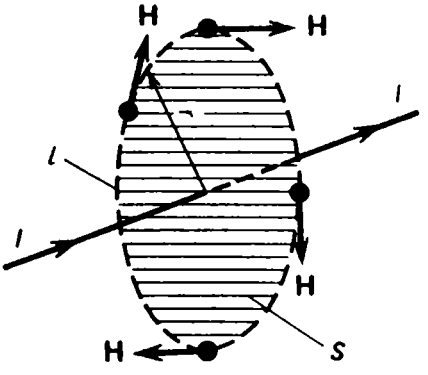


FIG. 3.20

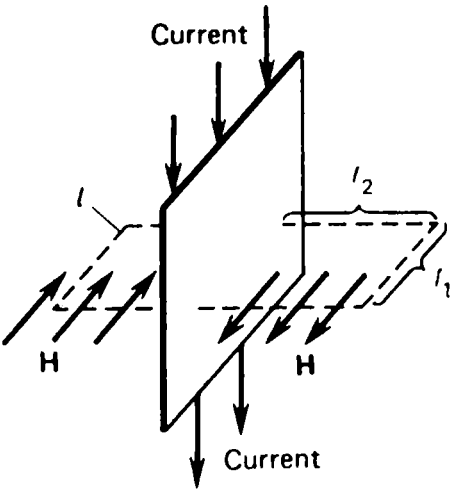


FIG. 3.21

tour is il_1. Hence (3.5.1) can be written as $H \cdot 2l_1 = (4\pi/c)il_1$. For the magnetic field of the current surface, we obtain the following expression:

$$H = \frac{2\pi}{c} i, \tag{3.5.3}$$

where i is the current per unit length in the direction perpendicular to the direction of motion of the charges.

3.5.3. Magnetic Moment and Its Relation with Mechanical (Angular) Momentum. Let us consider the motion of an isolated charge in the vicinity of an attracting charge centre. Let us assume that the trajectory of the charge is circular. Let us estimate the mean magnetic field produced by the mov-

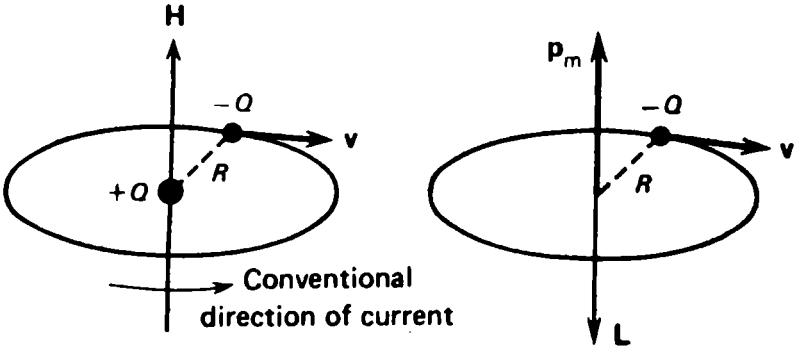


FIG. 3.22

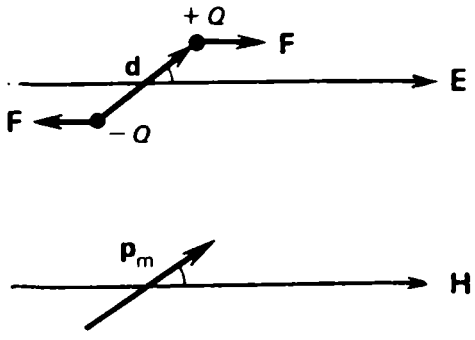


FIG. 3.23

ing charge at the centre of this circle. To calculate H, we use formula (3.5.2) for an infinite direct current with certain stipulations (a ring has neither beginning nor end). Naturally, in this case we must neglect the numerical factors, and hence the computation becomes just a rough approximation. Thus, we assume that $\langle H\rangle \approx (1/c)(I/R)$ at the centre of the circle. The current is given by $I = Q[v/(2\pi R)]$, where the multiplication factor for the charge is the number of revolutions per second. We introduce the quantity

$$\mathbf{p}_m = \frac{Q}{2c}[\mathbf{Rv}] \tag{3.5.4}$$

called the magnetic moment of the orbital motion of a charge (Fig. 3.22). We can thus write

$$\langle \mathbf{H}\rangle \approx \frac{\mathbf{p}_m}{R^3}\,. \tag{3.5.5}$$

Observe that the magnetic field is expressed in terms of the magnetic moment by the same formula as the expression for the electric field in terms of the (electric) dipole moment.

Let us establish a relation between the magnetic moment and angular momentum. From (3.5.4), we have

$$\mathbf{p}_m = \frac{Q}{2mc}[\mathbf{Rp}] = \frac{Q}{2mc}\mathbf{L}, \tag{3.5.6}$$

where L is the angular momentum of the particle whose velocity v is assumed to be much smaller than c.

If an electric dipole and a system having a magnetic moment are respectively placed in an electric and a magnetic field, the external field produces orienting moments. This situation is illustrated in Fig. 3.23.

3.6. Motion of Charges in a Field

The motion of charges in electromagnetic fields is investigated with the help of Eq. (3.2.2):

$$\frac{\Delta \mathbf{p}}{\Delta t} = Q\mathbf{E} + Q\left[\frac{\mathbf{v}}{c}\mathbf{H}\right].$$

Throughout this section, we shall assume that the charges move with a velocity $v \ll c$, and hence $\mathbf{p} = m\mathbf{v}$ is the ordinary particle momentum.

3.6.1. Motion of a Charge in a Constant Uniform Electric Field. Time-independent fields are called **constant fields.** Moreover, if the field strength is the same at all points in space, the field is called **uniform.** For a constant uniform electric field, the equation of motion of a charge has the form

$$\frac{\Delta \mathbf{p}}{\Delta t} = Q\mathbf{E}. \tag{3.6.1}$$

This equation is mathematically analogous to the equation

$$\frac{\Delta \mathbf{p}}{\Delta t} = m\mathbf{g},$$

describing the motion of a material particle in a gravitational field (see Sec. 2.4). Since similar equations have similar solutions, we can write by analogy

$$x = \frac{Q}{m}\left(\frac{E}{2v_0^2}\right)y^2. \tag{3.6.2}$$

We have thus arrived at an equation for the trajectory of a charge Q moving in an electric field E directed along the X-axis. The charged particle moves in the XY-plane, its mass and initial velocity being denoted by m and v_0 respectively.

In Sec. 2.4, we obtained the equation $x = [g/(2v_0^2)]y^2$ for the trajectory of a body moving in a uniform gravitational field. This equation has the same form for bodies of different masses since the expression for the trajectory does not contain mass. This is a fundamental property of the gravitational field. However, the trajectory (3.6.2) of a charge moving in an electric field depends on the ratio Q/m. This property of the electric field is used in the methods

of mass spectroscopy for determining the masses, say, of isotopes, viz. nuclei having the same electric charge but different masses.

3.6.2. Motion of a Charge in a Constant Uniform Magnetic Field. In this case, the equation of motion has the form

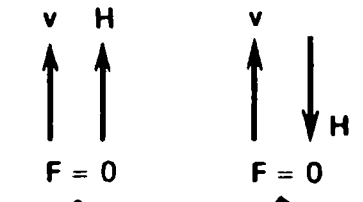


FIG. 3.24

$$\frac{\Delta \mathbf{p}}{\Delta t} = Q\left[\frac{\mathbf{v}}{c}\mathbf{H}\right]. \tag{3.6.3}$$

Since

$$|\mathbf{F}| = \frac{Q}{c}|\mathbf{v}|\,|\mathbf{H}|\sin(\widehat{\mathbf{vH}}) \tag{3.6.4}$$

in Eq. (3.6.3), the Lorentz force vanishes when a charge moves along or against the magnetic field (Fig. 3.24). If, however, the charge moves at right angles to **H**, the Lorentz force $\mathbf{F} = (Q/c)vH \neq 0$ (Fig. 3.25*a*).

This allows us to conclude (see Eq. (3.4.8)) that *the magnetic field does not perform any work on the electric charges.* In other words, the energy of charges in the field **H** is conserved. The motion of charges is circular in the case $\mathbf{v} \perp \mathbf{H}$ (Fig. 3.25*b*). For such a motion, the acceleration is constant and obtained as a combination of invariable kinematic quantities v^2 and R. The equation of motion (3.6.3) in this case can be written in the form

$$m\frac{v^2}{R} = \frac{Q}{c}vH. \tag{3.6.5}$$

Taking (2.6.29) into consideration, we obtain from (3.6.5)

$$\omega = \frac{QH}{mc}, \tag{3.6.6}$$

where ω is the rotational frequency of the charge in the magnetic field.

If the velocity **v** of the charge forms an angle α with the field **H** (Fig. 3.26*a*), the charge moves along a helical trajectory as shown in Fig. 3.26*b*.

3.6.3. Motion of a Charge in a Coulomb Field. A **Coulomb field** is produced by a stationary point charge Q_1, and the field strength $E = Q_1/r^2$ (see Eq. (3.4.1)) is a function of

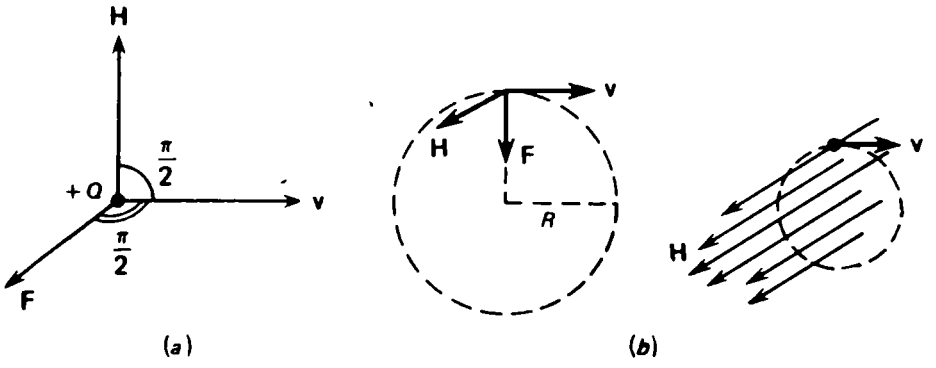


FIG. 3.25

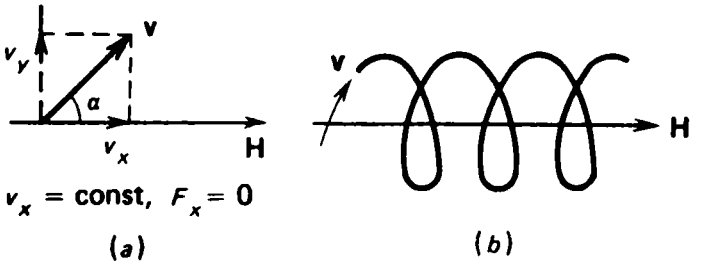


FIG. 3.26

only the distance r to the point of observation. Suppose that a charge Q_2 moves in such a field. Then the energy of interaction of these charges is $U = Q_1Q_2/r$ (see Eq. (3.4.10)). We assume that the mass of the positive charge Q_1 (centre) is large in comparison with the mass m of the moving negative charge Q_2.

A Coulomb field has a central symmetry, the centre coinciding with the position of the charge Q_1. All directions in space passing through this centre are equivalent. This means (see Sec. 2.6.3) that the angular momentum of the moving charge Q_2 is conserved. The charge moves in a plane (Fig. 3.27*a*) perpendicular to the angular momentum vector **L**.

It is convenient to consider the motion in **polar coordinates** (Fig. 3.27*b*). The polar coordinates r and φ are connected to the Cartesian coordinates x and y through the relations

$$x = r\cos\varphi, \quad y = r\sin\varphi. \tag{3.6.7}$$

Since the interaction energy U in the problem under consideration is independent of time, the total energy E of the system, viz. the sum of the kinetic energy T and the poten-

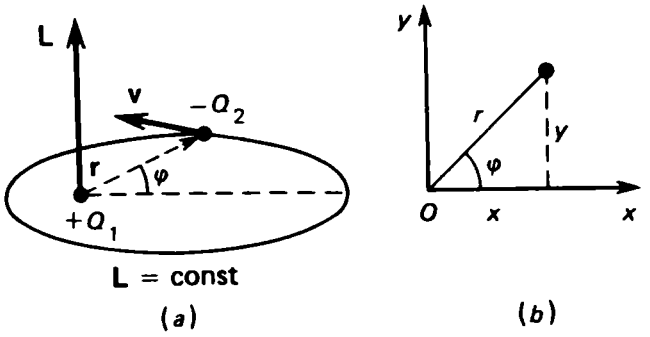


FIG. 3.27

tial energy U, must be conserved (see the end of Sec. 2.6.2). Let us express the kinetic energy

$$T = \frac{m}{2} v^2 = \frac{m}{2} (v_x^2 + v_y^2) \tag{3.6.8}$$

in terms of polar coordinates.

For this purpose, we derive an auxiliary mathematical relation for the product of two arbitrary functions $f(x)$ and $g(x)$:

$$\Delta (fg) = f(x + \Delta x)\, g(x + \Delta x) - f(x)\, g(x)$$
$$= (f(x) + \Delta f)\, (g(x) + \Delta g) - f(x)\, g(x)$$
$$= fg + g\,\Delta f + f\,\Delta g + \Delta f\,\Delta g - fg$$
$$= g\,\Delta f + f\,\Delta g + \text{small terms proportional to } (\Delta x)^2.$$

Finally, we obtain

$$\Delta (fg) = g\,\Delta f + f\,\Delta g. \tag{3.6.9}$$

Applying this relation to (3.6.7), we get

$$\Delta x = \Delta r \cos \varphi + r \Delta (\cos \varphi).$$

But since $\Delta (\cos \varphi) = -\sin \varphi\, \Delta \varphi$ (this equality can easily be obtained from Fig. 2.11), we find

$$\frac{\Delta x}{\Delta t} = \frac{\Delta r}{\Delta t} \cos \varphi - \frac{\Delta \varphi}{\Delta t} r \sin \varphi$$

or, in alternative form,

$$v_x = v_r \cos \varphi - \omega r \sin \varphi, \tag{3.6.10}$$

where $\omega = \Delta\varphi/\Delta t$ is the angular velocity. Carrying out analogous operations, we obtain the following expression for v_y:

$$v_y = v_r \sin \varphi + \omega r \cos \varphi. \tag{3.6.11}$$

Here, we have used the relation $\Delta (\sin \varphi) = \cos \varphi \, \Delta\varphi$ obtained in Sec. 2.6.3 (see Eq. (2.6.26)). Substituting (3.6.10) and (3.6.11) into (3.6.8), we finally arrive at the expression

$$T = \frac{m}{2} (v_r^2 + \omega^2 r^2). \tag{3.6.12}$$

Hence the total energy of the system can be written in the form

$$E = T + U = \frac{m}{2} (v_r^2 + \omega^2 r^2) - \frac{Q_1 Q_2}{r} . \tag{3.6.13}$$

For the sign of the potential energy, see Eq. (3.4.10) as well as the remarks concerning it and Fig. 3.18. The notation used for the potential energy term in Eq. (3.6.13) can be conveniently used for analysis in polar coordinates.

For studying the motion of a charge in a Coulomb field, we proceed from the conservation laws rather than from the equations of motion.

The expression for angular momentum in polar coordinates is obtained from Eq. (2.6.30):

$$L = m\omega r^2.$$

Using this relation in the expression for the total energy, we can present (3.6.13) in the following form:

$$E = \frac{m}{2} v_r^2 + \frac{L^2}{2mr^2} - \frac{Q_1 Q_2}{r} . \tag{3.6.14}$$

The last two terms on the right-hand side of this expression can be considered to be the effective potential energy:

$$U_{\text{eff}} = \frac{L^2}{2mr^2} - \frac{Q_1 Q_2}{r} . \tag{3.6.15}$$

Indeed, in view of the law of angular momentum conservation, we obtain the expression $U_{\text{eff}} = A/r^2 - B/r$ (A and B are constants) which is a function only of r. We represent U_{eff} graphically as shown in Fig. 3.28 and consider the mo-

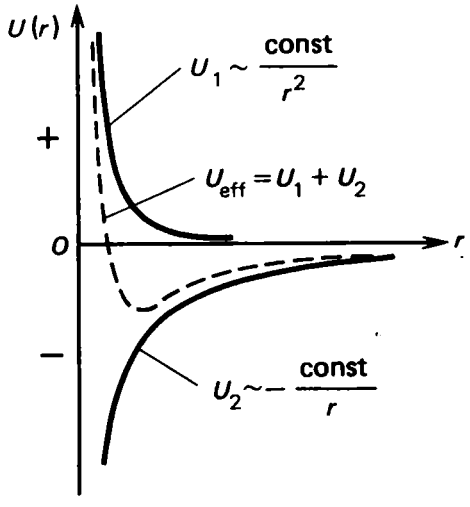


FIG. 3.28

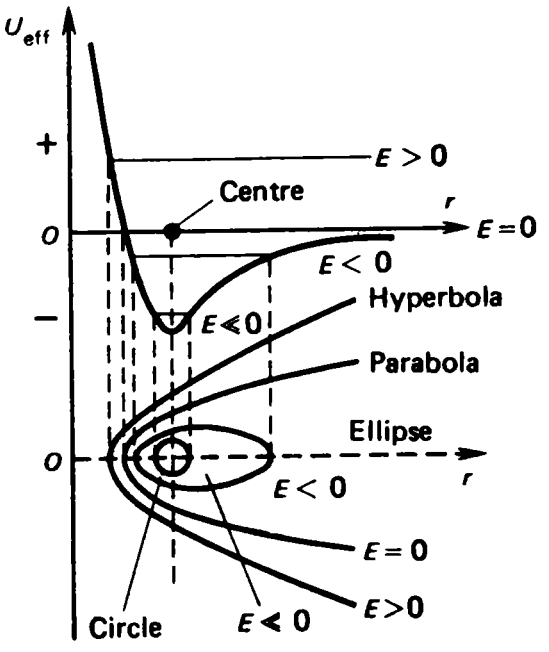


FIG. 3.29

tion of a charge in such a potential field (Fig. 3.29). The turning point on the trajectory corresponds to the condition $E = U_{\text{eff}}$. If the total energy E of the system is negative, the charge moves inside a "potential well", its possible trajectories being a circle and an ellipse. For $E \geqslant 0$, the motion is restricted only from one side. In other words, for positive values of U_{eff}, the system Q^+-Q^- of charges is no longer a bound system.

Since the mathematical form of the **law of universal gravitation**[6] $F = -Gm_1m_2/r^2$, where G is the **gravitational constant**, is analogous to expression (3.4.6) for Coulomb's law, all the results obtained above are automatically applicable to the motion of massive objects in a central gravitation-

[6] The potential energy of gravitational interaction of bodies is always negative (gravitation is manifested only as an attraction of bodies). In order to go over from the interaction forces F to the interaction energy U, we make use of the relation $F = -\mathrm{d}U/\mathrm{d}r$ between them. This gives

$$U = -\int F\,\mathrm{d}r = -\int \left(-\frac{Gm_1m_2}{r^2}\right)\mathrm{d}r = Gm_1m_2\int \frac{\mathrm{d}r}{r^2} = -\frac{Gm_1m_2}{r}.$$

As $r \to \infty$, we find that $U \to 0$, which means that the energy of bound systems is always negative. For freely moving objects, the energy $mv^2/2$ is always positive.

al field (Kepler's problem). The reader is surely familiar with the facts like the circular and elliptical orbits of the Earth's artificial satellites, orbital velocity, escape velocity.

3.7. Fields of Moving Charges. Emission

Let us study the effect of the motion of charges on the field created by them. In these problems, we shall not confine ourselves to the case $v \ll c$ so that the relativistic effects are fully manifested.

3.7.1. Field of a Uniformly Moving Charge. We are aware that the electric field of a stationary charge is an ordinary Coulomb field having a central spherical symmetry with the centre coinciding with the position of the charge. How does the electric field of the charge vary as a result of its motion with a constant velocity, and what is the nature of this variation? We shall solve this problem in several stages:

(i) knowing the field of a stationary charge (in the reference frame (X', Y')), we shall determine the field of a moving charge (in the reference frame (X, Y)); for this purpose, we shall naturally make use of the laws of field transformations;

(ii) the electric field of a moving charge must be expressed in terms of the coordinates of the reference frame relative to which the charge is moving (such a coordinate transformation must be carried out on the basis of the Lorentz transformations); this, however, may not be an easy task and requires some ingenuity (a significant role is played by diagrams in this case);

(iii) to study the configuration of the electric field of a fast charged particle, it is convenient to use angular coordinates.

Let us now consider the first stage in the solution of this problem. We place a charge Q at the origin of coordinates of the inertial reference frame (X', Y') moving relative to another inertial reference frame (X, Y) with a velocity $\mathbf{V}$ (Fig. 3.30*a*). The electric field $\mathbf{E}'$ in the reference frame (X', Y') is the ordinary Coulomb field of a stationary charge (Fig. 3.30*b*):

$$\mathbf{E}' = \frac{Q\mathbf{R}'}{(R')^3} . \qquad (3.7.1)$$

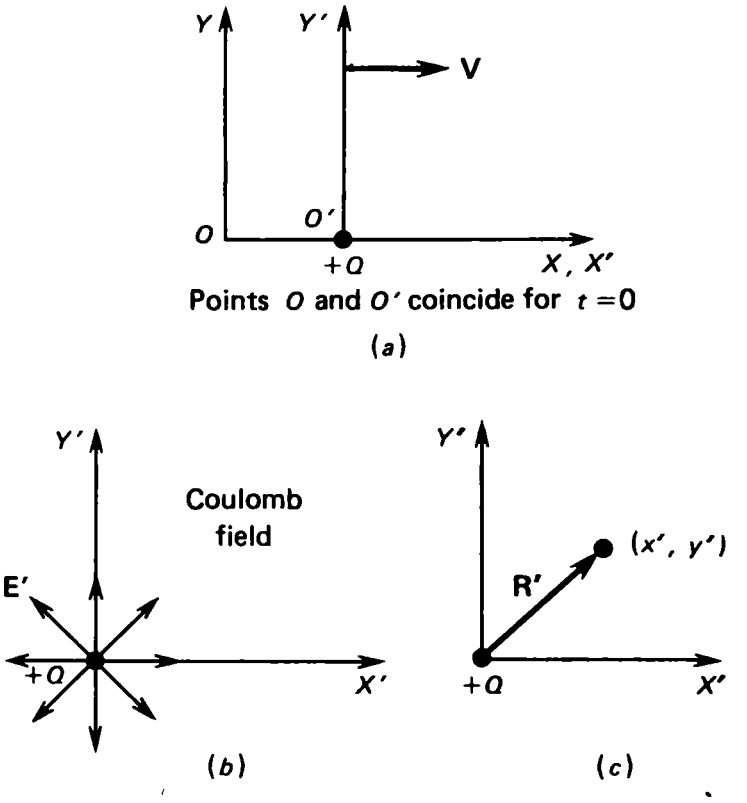


FIG. 3.30

This is the vector form of notation for Eq. (3.4.1), where we have introduced the unit vector $\mathbf{n} = \mathbf{R}'/R'$ in the direction of $\mathbf{R}'$ (Fig. 3.30*c*).

In accordance with the field transformation formulas (3.2.8), the components of the electric field in the reference frame (X, Y) can be written as follows:

$$E_x = E'_x = \frac{Qx'}{(R')^3}, \quad E_y = \frac{E'_y}{\sqrt{1-V^2/c^2}} = \frac{Qy'}{(R')^3\sqrt{1-V^2/c^2}}. \tag{3.7.2}$$

Here, we have taken into consideration the fact that $\mathbf{H}' = 0$ in the reference frame (X', Y') since the charge is at rest in this frame.

Next we consider the second stage in the solution of the problem. We express R' in terms of x and y with the help of the Lorentz transformations (2.7.1):

$$x' = \frac{x-Vt}{\sqrt{1-V^2/c^2}}, \quad y' = y. \tag{3.7.3}$$

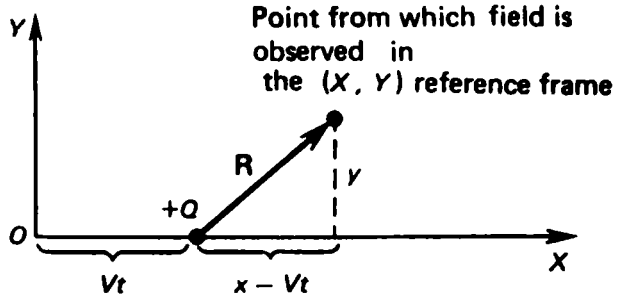


FIG. 3.31

These expressions are the inversions of transformations (2.7.1) since they can be obtained by substituting $-V$ for V into them. This is so because the reference frame (X, Y) moves with a velocity $-V$ relative to the reference frame (X', Y'). Consequently,

$$(R')^2=(x')^2+(y')^2=\frac{(x-Vt)^2+(1-V^2/c^2)\,y^2}{1-V^2/c^2}\,. \tag{3.7.4}$$

We now substitute the expressions for x', y', and R' in terms of x and y into (3.7.2). This gives

$$E_x=\frac{Q\,(x-Vt)\,(1-V^2/c^2)^{3/2}}{\sqrt{1-V^2/c^2}\,[(x-Vt)^2+(1-V^2/c^2)\,y^2]^{3/2}}\,,$$

$$E_y=\frac{Qy\,(1-V^2/c^2)^{3/2}}{\sqrt{1-V^2/c^2}\,[(x-Vt)^2+(1-V^2/c^2)\,y^2]^{3/2}}\,.$$

Hence

$$\left.\begin{aligned}E_x&=Q\left(1-\frac{V^2}{c^2}\right)\frac{x-Vt}{[(x-Vt)^2+(1-V^2/c^2)\,y^2]^{3/2}}\,,\\ E_y&=Q\left(1-\frac{V^2}{c^2}\right)\frac{y}{[(x-Vt)^2+(1-V^2/c^2)\,y^2]^{3/2}}\,.\end{aligned}\right\} \tag{3.7.5}$$

For a more visual presentation of the field components in these relations, we consider the structure of formulas for E_x and E_y and use the vector form of notation for the field $\mathbf{E}$. Here it is natural to introduce the vector $\mathbf{R}$ with the components $R_x = x - Vt$ and $R_y = y$ (Fig. 3.31), as well as the notation

$$(R^*)^2=(x-Vt)^2+\left(1-\frac{V^2}{c^2}\right)y^2. \tag{3.7.6}$$

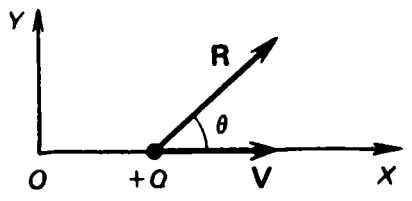


FIG. 3.32

Formulas (3.7.5) can then be written in the form

$$\mathbf{E} = \left(1 - \frac{V^2}{c^2}\right) \frac{Q\mathbf{R}}{(R^*)^3} . \qquad (3.7.7)$$

This is the final form of the expression for the electric field of a uniformly moving charge.

Finally, we turn to the last stage in the solution of this problem. To study the nature of the field (3.7.7), we introduce the angle θ between the direction of motion and the radius vector **R** (Fig. 3.32). In this case, $y^2 = R^2 \sin^2 \theta$, and Eq. (3.7.6) can be written in the form

$$(R^*)^2 = R_x^2 + \left(1 - \frac{V^2}{c^2}\right) R^2 \sin^2 \theta.$$

Since

$$R_x^2 = R^2 - R_y^2 = R^2 - y^2 = R^2 (1 - \sin^2 \theta),$$

we obtain

$$(R^*)^2 = R^2 \left(1 - \frac{V^2}{c^2} \sin^2 \theta\right). \qquad (3.7.8)$$

Substituting (3.7.8) into (3.7.7), we get

$$\mathbf{E} = \frac{Q\mathbf{R}}{R^3} \frac{1 - V^2/c^2}{[1 - (V^2/c^2) \sin^2 \theta]^{3/2}} . \qquad (3.7.9)$$

Let us analyze this relation. For $V \ll c$, this field is the same as the ordinary Coulomb field $\mathbf{E} = Q\mathbf{R}/R^3$ (see Eqs. (3.4.1) and (3.7.1)). Next, we note the distance R between the charge and the point of observation in the field and draw a sphere of radius R. Then the value of the field **E** will be minimum along the direction of motion ($\theta = 0, \pi$) and is given by

$$E_{\parallel} = \frac{Q}{R^2} \left(1 - \frac{V^2}{c^2}\right). \qquad (3.7.10)$$

The field has its highest value in the direction perpendicular to the velocity ($\theta = \pi/2$):

$$E_{\perp} = \frac{Q}{R^2} \frac{1}{\sqrt{1 - V^2/c^2}} . \qquad (3.7.11)$$

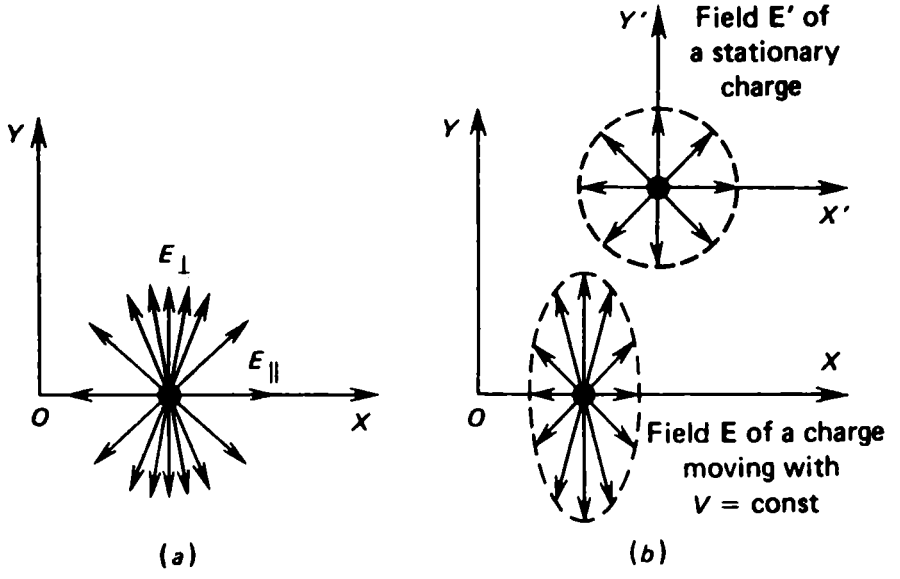


FIG. 3.33

It can be seen from (3.7.10) and (3.7.11) that as the velocity V of a moving charge increases, the field $E_\parallel$ decreases, while $E_\perp$ increases (Fig. 3.33). Two methods for a graphic representation of the electric field **E** are shown in Fig. 3.33. One of them connects the value of the field with the density of field lines (Fig. 3.33*a*), while in the other method, the magnitude of the field **E** is characterized by the length of the vector (Fig. 3.33*b*).

This behaviour of the electric field of a moving charge is manifested quite clearly for the velocities of motion comparable with the ultimate velocity c.

There is no magnetic field in the reference frame (X', Y'). The field is nonzero in the reference frame (X, Y) and is given, in accordance with (3.2.6), by

$$\mathbf{H} = \frac{1}{c}[\mathbf{VE}].$$

For $V \ll c$, we obtain

$$\mathbf{H} = \frac{1}{c}\left[\mathbf{V}\frac{Q\mathbf{R}}{R^3}\right] = \frac{Q}{c}\frac{[\mathbf{VR}]}{R^3}. \qquad (3.7.12)$$

This is the expression for the **Biot-Savart law.**

3.7.2. Emission by a Charge Moving with an Acceleration. At $t = 0$, a charge Q is at rest. It is then accelerated over a

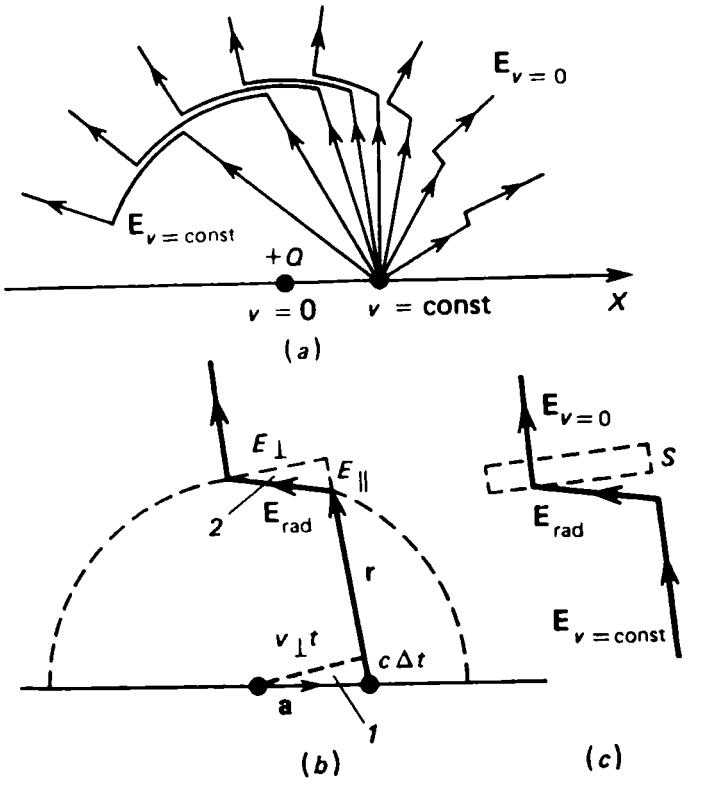


FIG. 3.34

period Δt with an acceleration **a** and continues to move with a constant velocity **v**. The electric field of such a moving charge is shown in Fig. 3.34*a*. The field lines $\mathbf{E}_{v=0}$ correspond to the field of the stationary point charge, while the lines $\mathbf{E}_{v=\text{const}}$ show the field of the charge moving with a constant velocity, and $\mathbf{E}_{\text{rad}}$ are the lines of the transient field represented in the form of a spherical "front" (see also Figs. 3.34*b* and *c*). The fields $\mathbf{E}_{v=0}$ and $\mathbf{E}_{v=\text{const}}$ are independent of time, while $\mathbf{E}_{\text{rad}}$ is variable. Like all varying fields, the field $\mathbf{E}_{\text{rad}}$ (called the **radiation field**) propagates in space with the velocity c. The radiation field has wave properties and has the form of a plane wave at large distances from the charge (see the next section).

Let us determine the intensity of the field $\mathbf{E}_{\text{rad}}$ at the point **r** at an instant of time t. This instant is chosen in such a way that the radiation wave passes through the point **r** at the instant t, i.e. $r = ct$. We assume that $t \gg \Delta t$, where Δt is the time for which the charge is accelerated, and that the charge velocity $v \ll c$. In this case, the displacement vt

of the charge is small in comparison with the distance ct traversed by the radiation field. Hence the lines of field $\mathbf{E}_{v=\text{const}}$ before inflection are assumed to be parallel to the lines of field $\mathbf{E}_{v=0}$ after inflection. As mentioned above, the inflection of the field lines is associated with the radiation field $\mathbf{E}_{\text{rad}}$. The geometrical constructions corresponding to this situation are represented in Fig. 3.34*b*. It can be seen from this figure that triangles *1* and *2* are congruent. This congruence leads to the relation

$$\frac{E_\perp}{E_\parallel} = \frac{v_\perp t}{c\,\Delta t}\,. \tag{3.7.13}$$

Here, $E_\perp$ and $E_\parallel$ are the components of the field $\mathbf{E}_{\text{rad}}$, $c\,\Delta t$ is the distance covered by the radiation wave in time Δt for which the charge is accelerated, and $v_\perp t$ is the displacement of the charge over a time t in the direction perpendicular to the direction of wave propagation. Since $v_\perp = a_\perp\,\Delta t$ and $t = r/c$, we can write (3.7.13) in the form

$$\frac{E_\perp}{E_\parallel} = \frac{(a_\perp\,\Delta t)\,(r/c)}{c\,\Delta t} = \frac{a_\perp r}{c^2}\,. \tag{3.7.14}$$

The component $E_\parallel$ of the radiation field can be obtained from the general field equations. We consider the first field equation (3.3.3) and construct an auxiliary closed surface S in the form of a "box" (Fig. 3.34*c*). Since there are no charges inside the box, the total flux of the electric field through it is zero. Consequently, the component $E_\parallel$ of the radiation field normal to the surface S must be equal to the component of the field $\mathbf{E}_{v=0}$ of the stationary charge normal to S, i.e. $E_\parallel = Q/r^2$ (see Eq. (3.4.1)). Substituting this expression for $E_\parallel$ into (3.7.14), we obtain

$$E_\perp = \frac{Q}{c^2}\,\frac{a_\perp}{r}\,. \tag{3.7.15}$$

Since the component $E_\parallel$ is independent of time, it cannot be included in the radiation wave. Thus, the radiation field is a transverse wave, i.e. the vector $\mathbf{E}_{\text{rad}}$ is perpendicular to the direction of wave propagation. The vector notation for (3.7.15) has the form

$$\mathbf{E}_{\text{rad}}(\mathbf{r},\ t) = -\frac{Q}{c^2}\,\frac{\mathbf{a}_\perp(t')}{r}\,. \tag{3.7.16}$$

Here, $t' = t - r/c$ is an instant of time earlier than t, and the minus sign indicates the opposite directions of the vectors $\mathbf{E}_{rad}$ and $\mathbf{a}$ (see Fig. 3.34*b*).

Let us obtain an expression for the energy of the radiation field. We proceed from the formula for the **energy flux density** $\mathbf{j}_E = \rho_E \mathbf{v}$, where ρ_E is the energy density, and $\mathbf{v}$ is the rate of energy transfer (cf. analogous expression for the current density in Eq. (3.3.5)). For the radiation field, $v = c$, and ρ_E is the energy of the field contained in a unit volume of space. Since we are considering the electric radiation field, the only quantity on which ρ_E may depend is $\mathbf{E}_{rad}$. Since ρ_E is a scalar, it can depend only on a scalar quantity. The scalar of $\mathbf{E}_{rad}$ is E^2_{rad}. Hence, with the exception of the numerical factors, the expression for the energy flux density of the radiation field can be written in the form

$$j_E \approx cE^2_{rad} \tag{3.7.17}$$

or, with the help of (3.7.16), in the form

$$j_E \approx \frac{Q^2}{c^3} \frac{a^2}{r^2} . \tag{3.7.18}$$

In order to estimate the total energy radiated per unit time, i.e. the radiation intensity J, we must multiply j_E by $4\pi r^2$. This follows from the fact that j_E is defined as the flux of energy passing through a unit area perpendicular to the direction of energy flux per second. Hence, omitting numerical factors, we obtain

$$J \approx \frac{Q^2}{c^3} a^2. \tag{3.7.19}$$

Formula (3.7.19) explains the mechanism of the emergence of electromagnetic bremsstrahlung in X-ray tubes (Fig. 3.35). A beam of high-energy electrons impinging on a tungsten anode in the tube is suddenly decelerated. This deceleration is accompanied by the emission of electromagnetic waves of very high frequency, with a fairly high penetrability.

3.7.3. Emission by a Charge Moving Uniformly in a Circle. The motion of a charge (say, electron) in a constant uniform magnetic field $\mathbf{H}$ is circular when $\mathbf{v} \perp \mathbf{H}$. In accordance with (3.6.5), the acceleration in this case is given by

$$a = \frac{eH}{mc} v.$$

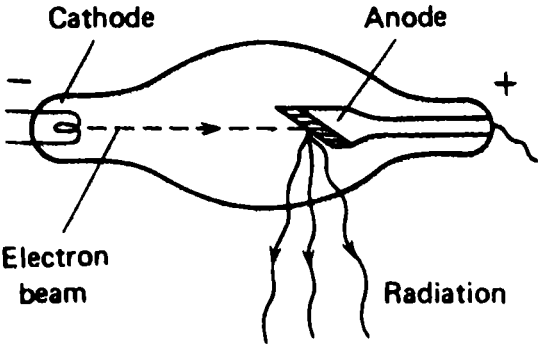


FIG. 3.35

Substituting this expression for centripetal acceleration into Eq. (3.7.19), we obtain the following expression for the radiation intensity:

$$J \approx \left(\frac{e^4}{c^5 m^2}\right) H^2 v^2. \tag{3.7.20}$$

This formula for the intensity of electromagnetic radiation emitted by charges moving in magnetic fields has been obtained for the case $v \ll c$. The radiation frequency mainly coincides with the rotational frequency (3.6.6) of charges around the lines of force for the field **H**.

By way of illustration, we can mention the cosmic radiation emitted by galaxies (motion of electrons in the magnetic fields of stars) and the electromagnetic radiation emitted by electrons in betatron-type accelerators (it should be noted that the radiation may fall in the visible spectral region, when the "luminous electron effect" is observed).

Let us consider the model of an atom in which an electron moves around a proton in accordance with the laws of classical mechanics. In view of the loss of energy due to radiation, the electron falls into the centre in a very short time, and the atom as such ceases to exist. The lifetime of such an "atom" can be estimated as follows. According to (3.4.10), the initial electrostatic energy of the "electron-proton" system is $U_0 \sim e^2/R_0$. By definition, the radiation intensity $J \sim U_0/\Delta t$, where we use formula (3.7.19) for J, and the acceleration is given by $a = v^2/R_0$. Consequently, $\Delta t \sim c^3 R_0/v^4$. Assuming that $R_0 \sim 10^{-10}$ m and $v \sim 10^6$ m/s, we obtain $\Delta t \sim 10^{-9}$ s.

3.8. Electromagnetic Waves

3.8.1. Some Properties of Radiation Fields. It was mentioned above that the field $\mathbf{E}_{rad}$ is perpendicular to the direction of field propagation. It follows from the law (3.2.6) of magnetic field transformation,

$$\mathbf{H} = \frac{1}{c}[\mathbf{VE}],$$

that for $\mathbf{V} = \mathbf{c}$

$$\mathbf{H}_{rad} = \left[\frac{\mathbf{c}}{c}\mathbf{E}_{rad}\right] = [\mathbf{n}_c\mathbf{E}_{rad}],$$

where $\mathbf{n}_c$ is a unit vector in the direction of radiation field propagation. As a result, we obtain the following relations for the radiation field:

$$|\mathbf{H}_{rad}| = |\mathbf{E}_{rad}|$$

and

$$\mathbf{H}_{rad} \perp \mathbf{E}_{rad} \perp \mathbf{n}_c. \qquad (3.8.1)$$

The mutual orientation of these three vectors is shown in Fig. 3.36.

3.8.2. Travelling Waves. Suppose that the vectors **E** and **H** are oriented in the radiation field in such a way that the direction of propagation $\mathbf{n}_c$ of the radiation coincides with the X-axis (Fig. 3.37). In this case, if the field characteristics at the initial instant of time $t = 0$ had a certain value at the point x in space, they will have the same value after a time t at a point separated by a distance ct from the initial point along the positive direction of the X-axis. In other words, the field characteristics are the same for x-coordinates and time instants t satisfying the relation

$$t - \frac{x}{c} = \text{const}.$$

Consequently, the radiation field forms a plane travelling above. In this case, the electromagnetic wave is transverse, i.e. the vectors **E** and **H** oscillate in a plane perpendicular

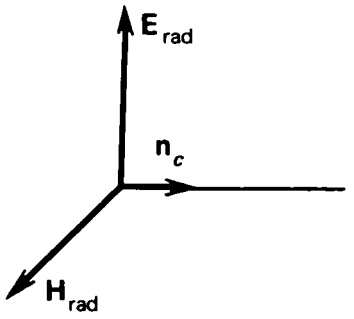


FIG. 3.36

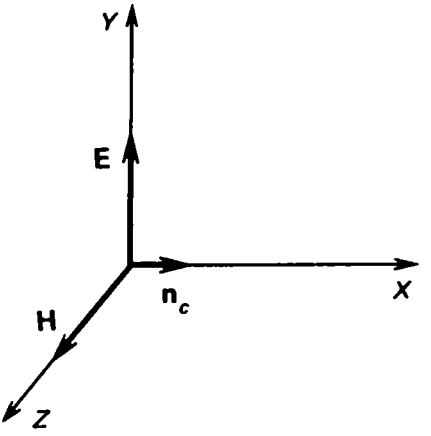


FIG. 3.37

to the direction of wave propagation (Fig. 3.38). The components E_y and H_z of the field in an electromagnetic wave are simple periodic functions of $(t - x/c)$:

$$\left.\begin{aligned} E_y &= A\cos\omega\left(t-\frac{x}{c}\right), \\ H_z &= A\cos\omega\left(t-\frac{x}{c}\right). \end{aligned}\right\} \qquad (3.8.2)$$

Such a wave is called a **monochromatic wave**. In (3.8.2), A is the **amplitude**, and ω is the **cyclic frequency**. The meaning of these concepts can clearly be understood by considering the example of mechanical vibrations.

The oscillations of a pendulum were considered in Sec. 2.6.3. The kinematics and the method of describing these oscillations are shown in Fig. 3.39. It can be seen from the figure that the displacement of a material particle from the

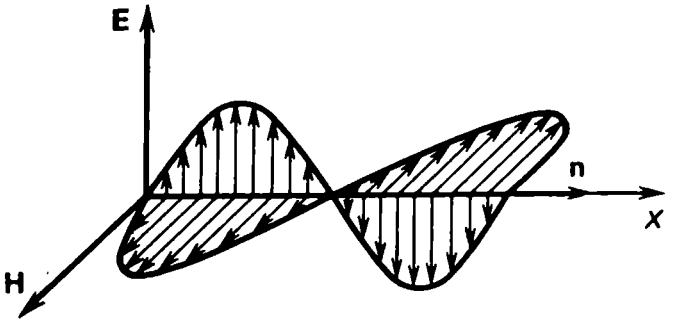


FIG. 3.38

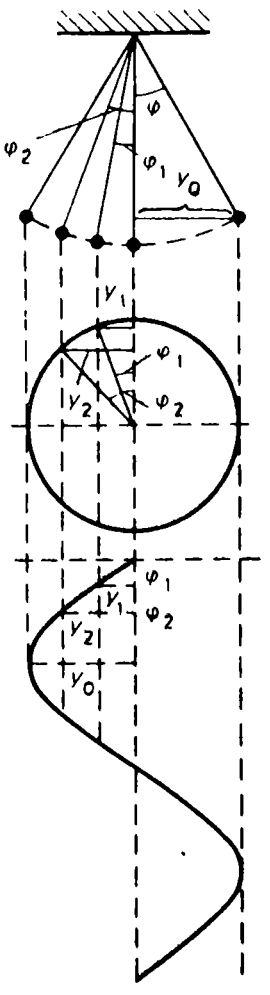


FIG. 3.39

equilibrium position is given by

$$y = y_0 \sin \varphi, \tag{3.8.3}$$

where y_0 is the maximum displacement (amplitude), and φ is the **phase**. Here,

$$\varphi = \omega t, \tag{3.8.4}$$

where ω is also the cyclic frequency. Denoting by T the **period of one oscillation** and considering that $T = 2\pi$ for a sine function, we obtain

$$\omega = \frac{2\pi}{T}. \tag{3.8.5}$$

When considering the propagation of waves, the concept of **wavelength** λ is introduced. This quantity is expressed in terms of T and the velocity of propagation c (for electromagnetic waves) as follows:

$$\lambda = cT,$$

or

$$\lambda = \frac{2\pi c}{\omega}. \tag{3.8.6}$$

3.8.3. Emission of Electromagnetic Waves by Oscillating Charges. Energy and Momentum of Waves. For a charge oscillating according to the law

$$y = y_0 \sin \omega t,$$

the velocity is given by

$$v_y = \frac{\Delta y}{\Delta t} = y_0 \omega \cos \omega t$$

and the acceleration is

$$a_y = \frac{\Delta v_y}{\Delta t} = -y_0 \omega^2 \sin \omega t$$

(see the derivation of formula (3.6.9)). As in any accelerated motion of a charge, the radiation intensity of electromagnetic waves is determined by (3.7.19), i.e.

$$J \approx \frac{Q^2}{c^3} y_0^2 \omega^4 \sin^2 \omega t.$$

Let us average this quantity over a time $t \geqslant T$. For this purpose, we use the trigonometric relation $\sin^2 \omega t = (1/2)(1 - \cos 2\omega t)$ whose mean value is equal to 1/2. Omitting the numerical factors, we obtain

$$J \approx \frac{Q^2}{c^3} y_0^2 \omega^4. \tag{3.8.7}$$

This relation indicates that *the radiation intensity of electromagnetic waves increases sharply with the frequency of oscillating electric charges.* Hence the frequency range used in practice begins from about 10^6 Hz (1 Hz means one oscillation per second).

Let us consider the energy and momentum of electromagnetic waves. It was shown in the previous section (see the remarks accompanying Eq. (3.7.17)) that the energy density ρ_E of an electric field is a quadratic function of the field **E**. For an electromagnetic field which is a superposition of electric and magnetic fields, the expression for ρ_E must be supplemented by the magnetic field **H**. Since **E** and **H** have the same dimensions (this follows from the expression (3.2.3) for the Lorentz force), we can write

$$\rho_E \approx E^2 + H^2. \tag{3.8.8}$$

The momentum density ρ_p of an electromagnetic wave can depend only on **E**, **H**, and c. Since ρ_p and ρ_E have dimensions $[\rho_p]$ = momentum/m^3 and $[\rho_E]$ = energy/m^3 respectively, it can easily be verified that the energy density and momentum density of an electromagnetic wave are connected through the relation

$$\rho_p = \frac{\rho_E}{c}. \tag{3.8.9}$$

The expression for the momentum density can be obtained by combining (3.8.8) and (3.8.9):

$$\rho_p \approx \frac{E^2 + H^2}{c}. \tag{3.8.10}$$

It should be observed that the momentum density is a vector directed along the direction n of wave propagation, i.e. $\rho_p = (\rho_E/c)\,\mathbf{n}$, where n is a unit vector.

3.8.4. Free Oscillations of a Field. Standing Waves. Travelling waves (3.8.2) propagate all over the space. Let us consider an electromagnetic wave enclosed in a box with mirror walls. The field vanishes at the boundaries of the box. There is no energy transfer, and the wave is called a **standing wave**. The electric field of such a wave has the form

$$E_y = A\,(x)\cos(\omega t + \alpha), \tag{3.8.11}$$

where α is the initial phase. We can present this equation in a different form. For this purpose, we make use of the trigonometric relation

$$\cos(\omega t + \alpha) = \cos\alpha\cos\omega t - \sin\alpha\sin\omega t.$$

Consequently, we get

$$E_y = C_1\cos\left(2\pi\frac{x}{\lambda}\right)\cos\omega t - C_2\sin\left(2\pi\frac{x}{\lambda}\right)\sin\omega t. \tag{3.8.12}$$

This form of notation indicates that *the standing wave is the result of interference of an incident wave and a reflected wave.*

We take into consideration the boundary condition $E_y = 0$ in (3.8.12) (Fig. 3.40):

$$0 = C_1\cos\omega t \quad \text{and} \quad C_1 = 0 \quad \text{for} \quad x = 0,$$

$$0 = -C_2\sin\left(2\pi\frac{L}{\lambda}\right)\sin\omega t \quad \text{for} \quad x = L,$$

i.e.

$$\sin\left(2\pi\frac{L}{\lambda}\right) = 0. \tag{3.8.13}$$

This equation has the following solution:

$$2\pi\frac{L}{\lambda} = n\pi,$$

where $n = 1, 2, 3, \ldots$. Alternatively, we can write

$$L = n\frac{\lambda}{2}, \tag{3.8.14}$$

where L represents the linear dimensions of the cavity (see Fig. 3.40).

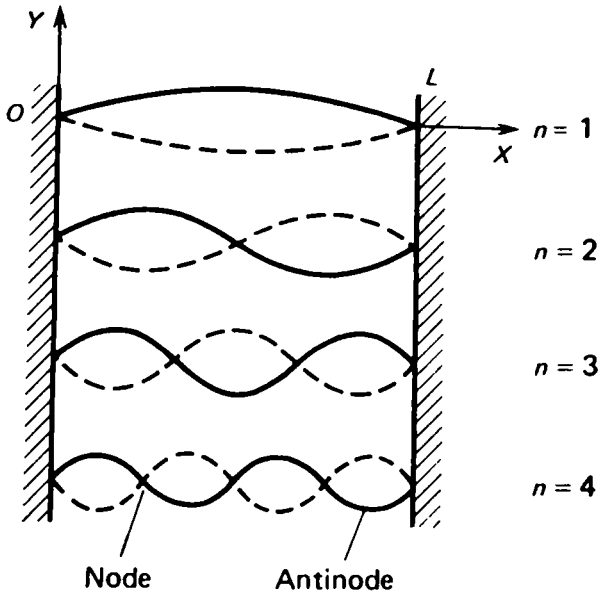


FIG. 3.40

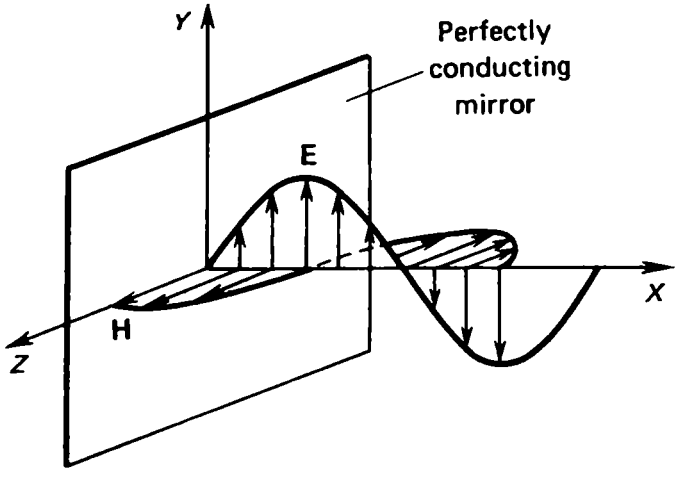


FIG. 3.41

Thus, the following states of the field are formed in a bounded region of space:

$$E_y(x,\ t) = C_2 \sin\left(2\pi \frac{x}{\lambda}\right) \sin \omega t \qquad (3.8.15)$$

only with a discrete set of frequencies

$$\omega = n\pi \frac{c}{L} . \qquad (3.8.16)$$

This expression for the spectrum of free field oscillations can be obtained from (3.8.14) with the help of (3.8.6).

Some other properties of standing electromagnetic waves are illustrated in Fig. 3.41.

3.9. Propagation of Light

The velocity of propagation of light was found to be equal to the velocity c of propagation of electromagnetic perturbations. The whole complex of experimental results indicates that light is nothing but electromagnetic waves in the wavelength region (4-8) $\times 10^{-7}$ m. Hence all optical phenomena must be considered within the framework of electrodynamics.

3.9.1. Interference of Electromagnetic Waves. Let us consider two waves f_1 and f_2 with the same frequency ω:

$$f_1 = A_1 \cos(\omega t + \alpha_1), \quad f_2 = A_2 \cos(\omega t + \alpha_2). \qquad (3.9.1)$$

Suppose that these waves are propagating in the same direction. Here, by f_1 and f_2 we mean either of the field components, E or H. For the sake of simplicity, we assume that they have the same amplitude ($A_1 = A_2 = A$). Then the resultant wave

$$f_R = f_1 + f_2,$$

obtained by superimposing two solitary waves, is given by

$$f_R = A\,[\cos(\omega t + \alpha_1) + \cos(\omega t + \alpha_2)].$$

Let us transform this expression. For this purpose, we make use of the trigonometric relation

$$\cos\alpha + \cos\beta = 2\cos\left(\frac{\alpha+\beta}{2}\right)\cos\left(\frac{\alpha-\beta}{2}\right).$$

Consequently, we obtain

$$f_R = 2A\cos\left(\frac{\alpha_1-\alpha_2}{2}\right)\cos\left(\omega t + \frac{\alpha_1+\alpha_2}{2}\right).$$

Thus, the resultant wave is also a cosine wave but with a new phase and a different amplitude:

$$A_R = 2A\cos\left(\frac{\alpha_1-\alpha_2}{2}\right).$$

Let us find the intensity (square of the amplitude) of the resultant wave:

$$A_R^2 = 4A^2 \cos^2 \left(\frac{\alpha_1 - \alpha_2}{2} \right).$$

Using the trigonometric relation

$$\cos^2 \alpha = \frac{1 + \cos 2\alpha}{2},$$

we obtain

$$A_R^2 = 2A^2 + 2A^2 \cos (\alpha_1 - \alpha_2). \qquad (3.9.2)$$

Going over to the general case of waves with different amplitudes (3.9.1), we can write (3.9.2) in the form

$$A_R^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos (\alpha_1 - \alpha_2). \qquad (3.9.3)$$

It can easily be verified that (3.9.2) is obtained from (3.9.3) for $A_1 = A_2 = A$. The cross-multiplication term in (3.9.3) containing the cosine term is called the **interference term** and may be positive or negative depending on the initial phase difference $(\alpha_1 - \alpha_2)$.

Let us use Eq. (3.9.2) to explain the interference effect. If the phase difference $\alpha_1 - \alpha_2 = \pi$, then $A_R^2 = 2A^2 - 2A^2 = 0$, while for $\alpha_1 - \alpha_2 = 0$, we get $A_R^2 = 2A^2 + 2A^2 = 4A^2$.

As a result of superposition, the waves may completely suppress each other. In the case of augmentation, the intensity of the resultant wave is different from the sum of the intensities of the component waves.

3.9.2. Diffraction of Electromagnetic Waves. The bending of waves round any kind of obstacle or a change in the direction of propagation of a wave passing, say, through a hole in a screen is called the **diffraction of waves**. Diffraction is especially noticeable when the size of the obstacle or the hole in the screen is comparable with the wavelength.

Let us consider diffraction at two slits (Fig. 3.42*a*) separated by a distance d of the order of the wavelength λ of the wave source (Fig. 3.42*b*). Waves from two slits *1* and *2* incident at any point P interfere with each other. Here, it is

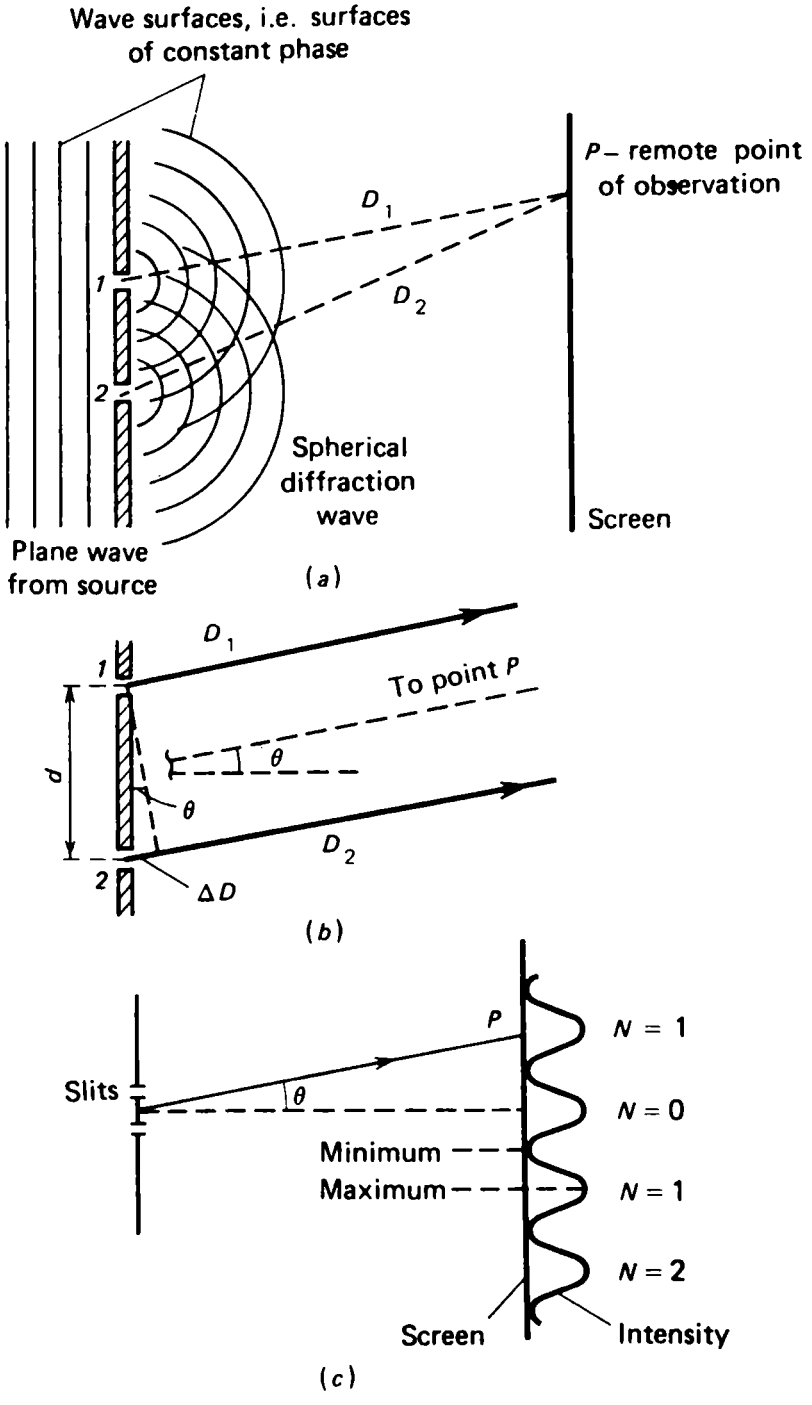


FIG. 3.42

convenient to go over from the phase difference $(\varphi_2 - \varphi_1)$ to the distance $D_2 - D_1 = \Delta D$ in the paths traversed by the waves. These quantities are connected through a very simple

relation: λ is the scale of length, and the ratio $\Delta D/\lambda$ multiplied by 2π is nothing but the phase difference. Consequently,

$$\varphi_2 - \varphi_1 = 2\pi \frac{d}{\lambda} \sin \theta, \tag{3.9.4}$$

where θ is the angle determining the direction at the point of observation. Putting $\varphi_2 - \varphi_1 = \pi$ in (3.9.4), which corresponds to the condition of the minimum intensity of the resultant wave, we arrive at the relation $d \sin \theta = \lambda/2$. In the general case, the **minimum intensity condition** is written in the form

$$d \sin \theta = (2N + 1) \frac{\lambda}{2}, \tag{3.9.5}$$

where $N = 0, 1, 2, \ldots$. Obviously, the **maximum intensity condition** can be presented in the form

$$d \sin \theta = 2N \frac{\lambda}{2}. \tag{3.9.6}$$

The interference pattern obtained as a result of diffraction at two slits is represented in Fig. 3.42*c*.

3.9.3. Geometrical Optics. The wavelength λ of light is usually much smaller than the size d of the bodies with which light interacts. Under these conditions, the wave properties of light (interference and diffraction) cannot be manifested. Indeed, if $\lambda \ll d$, it follows from (3.9.4) that the phase difference will be quite large. This means that the interference term in (3.9.3) will be practically zero since the cosine function for very large values of angles oscillates rapidly, and its mean value is zero. As far as diffraction is concerned, it follows from (3.9.6) written in the form $(d/\lambda) \sin \theta = N$ that the factor $\sin \theta$ must be close to zero for very large values of d/λ. In other words, light must propagate rectilinearly in this case.

The last result means that in the limiting case when $\lambda \to 0$, the concept of **rays** as "trajectories" of light is suitable for describing the laws of propagation of light.

Chapter 4
ATOMIC PHYSICS AND QUANTUM MECHANICS [1]

4.1. Planetary Model of Atom

From investigations of physical phenomena over large regions in space, we go over to extremely small regions (of the order of 10^{-10} m and smaller). Phenomena occurring in very small regions in space and involving particles having very small masses are studied in **atomic physics.**

If we try to apply the concepts of classical mechanics and electrodynamics to analyze the model of an atom as a system consisting of a positively charged nucleus containing practically the entire mass of the atom, and electrons revolving around the nucleus, we at once obtain results that are in sharp disagreement with the reality. Indeed, the motion of electrons in closed trajectories is obviously an accelerated motion and, from the point of view of classical theory, must be accompanied by the emission of electromagnetic waves. Thus, the electrons should continuously lose energy and ultimately fall into the nucleus. Estimates show (see the end of Sec. 3.7.3) that such an "atom" would have a lifetime of about 10^{-9} s. The very fact that macroscopic bodies exist in nature indicates that the particles constituting them, viz. atoms, have a stable structure.

This means that the intra-atomic motion must be considered by using new principles. The nature of these principles is revealed in experiments.

4.2. Experiments on Diffraction of Particles

Let us consider the experiments involving the diffraction of electrons by crystals. By passing solitary electrons in succession through a thin crystal (foil), scientists discov-

[1] CGS units are used in this chapter.

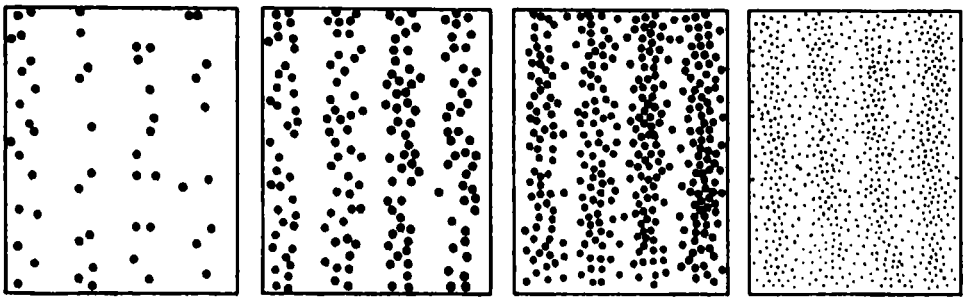


FIG. 4.1

ered a pattern of alternating maxima and minima of intensity on a screen (photographic plate).

In these experiments, the initial conditions and the conditions under which the electrons passed through the crystal were kept strictly identical. However, the electrons emerging through the crystal left tiny dark spots at different places on the photographic plate. The incidence of an ever larger number of solitary electrons on the screen led to the emergence of a definite diffraction pattern resulting from the distribution of the incident particles (Fig. 4.1). This experimental result indicates that although electrons have corpuscular properties, they behave like waves under the conditions mentioned above.

It is well known that the diffraction of any waves takes place provided that the wavelength is commensurate with the size of the obstacle or with the region of motion in general (for example, for electromagnetic waves, this is the size of the aperture in opaque screens; for the waves that can possibly be attributed to electrons, this range is the interatomic spacing in a crystal lattice).

According to the classical concepts concerning motion, a particle occupies a definite position in space at each instant of time and has a definite velocity. In classical mechanics, this means that the particle has a definite trajectory when it moves relative to large regions in space. This statement is valid to a considerable degree of accuracy for electrons travelling from an electron gun to a crystal. The situation changes drastically if the electrons move under the conditions in which the characteristic spatial region of motion is comparable with their wavelength. Here the wave effects

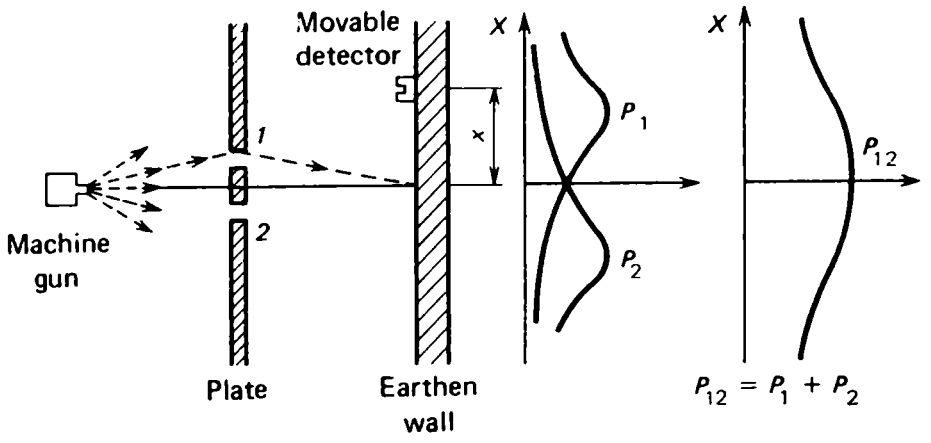


FIG. 4.2

assume prime importance, and the concept of trajectory for an electron as a particle is not valid in this case.

In order to present the unusual situation arising here in a more graphic form, let us successively consider a series of "mental experiments" involving bullets, waves, and electrons.[2]

Suppose that shots are fired at random from a machine gun (Fig. 4.2). An armoured plate with two holes *1* and *2* is placed in the path of the bullets. An earthen wall is situated behind the armoured plate, and a detector in the form of a sandbox is moved along this wall. The bullet is a body which is known to have a trajectory beforehand. This means that if we knew the initial conditions (coordinates and velocity), we could predict the motion of the bullet at all subsequent instants of time. However, since the firing is not carried out in a definite direction, the conditions under which the bullet passes through a hole change at random (for example, the bullet may ricochet from the edge of a hole). Hence the bullets land at quite different spots on the earthen wall.

With the help of a movable sandbox, we can count the number of bullets landing at different points on the earthen wall. We introduce the quantity $P = n/N$, where N is the total number of bullets fired from the machine gun in a certain interval of time, and n is the number of shots registered during this time in the detector with x-coordinates. Let us

[2] The idea of these experiments is borrowed from Feynman. See R. Feynman, *Theory of Fundamental Processes*, 1961.

plot the function $P(x)$. Naturally, a certain amount of smoothing will be required in view of the abrupt variation of P at adjacent sites. As a result, we obtain a probability distribution curve for a bullet to land at a certain spot on the earthen wall. After all, we cannot assign beforehand any specific position for a particular bullet.

The experiment involving the machine-gun firing is carried out as follows. We close hole *2* and leave hole *1* open. After firing and a series of measurements, we plot the probability distribution curve P_1. We then close hole *1* and open hole *2*. Once again, after firing and a quantitative analysis, we plot the probability distribution curve P_2. Finally, we leave both holes open. The curve obtained for the probability distribution in this case is indicated by P_{12} (see Fig. 4.2).

Irrespective of the position of the point where a bullet lands on the earthen wall, each bullet has its own trajectory. It follows hence that the probability distribution curve P_{12} can be presented just as the sum of the probability curves P_1 and P_2:

$$P_{12} = P_1 + P_2.$$

We now go over to the experiment with electromagnetic waves. We create conditions for the passage of a wave identical to those in the experiment with bullets (Fig. 4.3). Suppose that a plane wave is incident from an emitter onto a screen with two holes (the emitter is situated at a con-

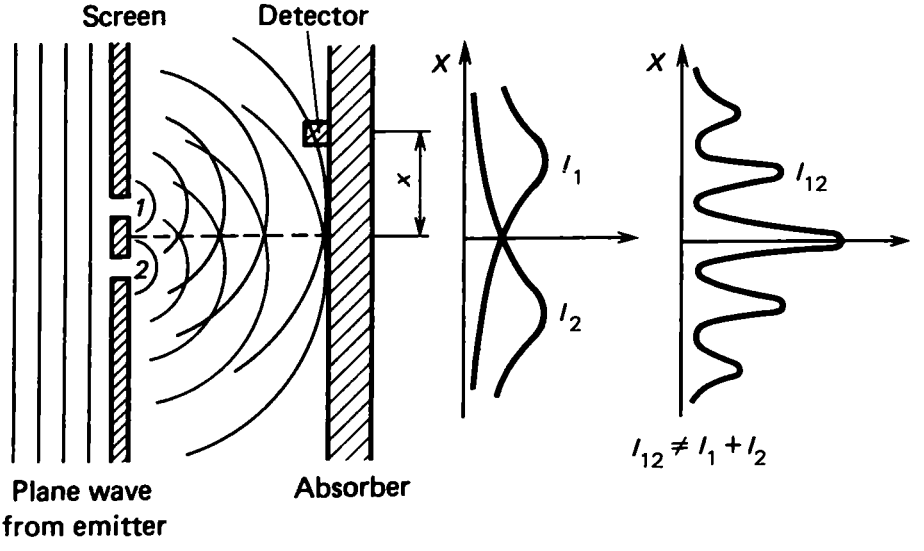


FIG. 4.3

siderable distance from the screen). We also assume that the distance between the holes and their size are comparable with the wavelength. In this case, a wave passing through a hole is diffracted and is incident on an absorbing screen with a detector (see Fig. 4.3). The detector may be in the form of a miniature antenna connected to a receiver of electromagnetic waves. This detector can register the wave intensity (defined as the square of the field strength), say, from the amplitude of the radio signal.

Leaving hole *1* open (and *2* closed), we obtain a certain distribution of the intensity I_1 of the wave field on the absorbing screen. Closing hole *1* and opening hole *2*, we can use the movable antenna to determine the intensity I_2 as a function of the x-coordinate of the antenna. When both holes are open, the waves passing through them interfere (see Sec. 3.9 and Fig. 3.42). As a result, we obtain the intensity

$$I_{12} \neq I_1 + I_2$$

(see formula (3.9.3)).

Finally, let us go over to experiments involving electrons. These experiments are a sort of idealization of the experiments on the diffraction of electrons by crystals. Suppose that the electrons from an electron gun are incident on a plate with two holes (Fig. 4.4). The separation between the holes is quite small (the criterion for the smallness of

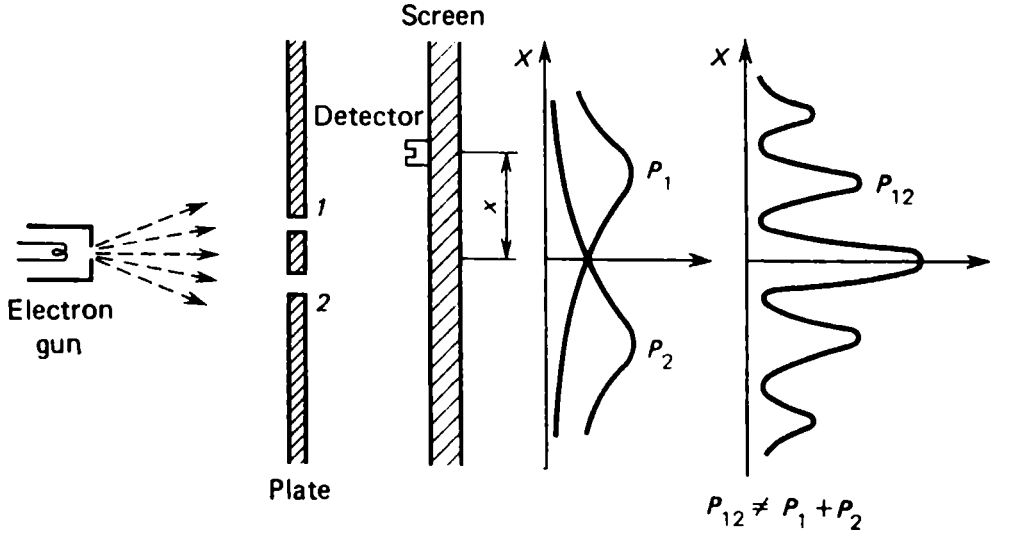


FIG. 4.4

the gap will be indicated below). A screen with a movable detector, say, a Geiger counter, is arranged behind this plate. The detector is capable of registering solitary electrons. Let us carry out the familiar operations. We close hole *2* and leave hole *1* open. Moving the detector along the screen, we register the incident electrons. Like bullets, the electrons are incident on the screen at different places in accordance with the laws of probability. As a result, we obtain a certain curve for the distribution of electrons over different points on the screen, i.e. the function $P_1(x)$. Then we open hole *2* (and close hole *1*) and once again follow the same procedure. This gives the curve $P_2(x)$. The whole set-up is just the same as in the case of bullets.

However, a strange thing happens if both holes are left open. In this case, the probability distribution curve P_{12} will be similar in form to the intensity distribution curve I_{12} in the experiment involving waves, i.e. in the case of electrons

$$P_{12} \neq P_1 + P_2.$$

In other words, interference effects are observed in experiments with electrons.

4.3. The Uncertainty Relation

If electrons participate in the interference phenomena, does it mean that they can annihilate one another? Certainly not! Each electron produces its own single dark spot on the photographic plate, and a fraction of an electron is never registered. The shape of the curve P_{12} for electrons merely indicates that the probability distribution for electrons appears at a certain point on the photographic plate.

With the help of the usual procedure (see Sec. 3.9), we can obtain from the diffraction experiments involving electrons the wavelength λ corresponding to the electron having a momentum p. Naturally, there must be a link between the wave and corpuscular properties of electrons. This connection is expressed through the relation

$$\lambda = 2\pi \frac{\hbar}{p}, \tag{4.3.1}$$

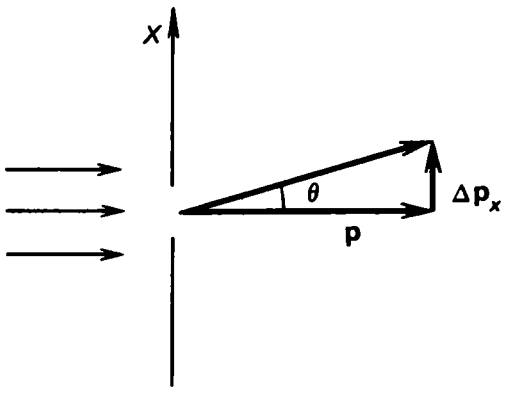


FIG. 4.5

where $\hbar = 1.05 \times 10^{-34}$ J·s, called **Planck's constant**, is a fundamental constant characterizing phenomena on the atomic scale.

Let us quantitatively analyze the diffraction of electrons by a crystal. Let Δx be the lattice constant (slit width), and θ the diffraction angle by which the electrons are deflected from their original direction of motion. The condition for the first maximum in the diffraction pattern is given (see Eq. (3.9.6) and Fig. 3.42) by the expression

$$\frac{\lambda}{\Delta x} \sim \sin\theta$$

or, considering that the angles are close to 90°,

$$\frac{\lambda}{\Delta x} \sim 1. \qquad (*)$$

Carrying out a similar analysis for the change Δp_x in the component of the electron momentum due to diffraction (Fig. 4.5), we obtain

$$\frac{\Delta p_x}{p} \sim 1, \qquad (**)$$

where p is the initial momentum of an electron in the beam. Combining (*) and (**), we obtain

$$\frac{\lambda}{\Delta x} \sim \frac{\Delta p_x}{p},$$

or

$$\Delta p_x \, \Delta x \sim \lambda p. \qquad (***)$$

Finally, taking into consideration Eq. (4.3.1), we can present (***) in the form

$$\Delta p_x \, \Delta x \geqslant \hbar \qquad (4.3.2)$$

and, similarly,

$$\Delta p_y \, \Delta y \gtrsim \hbar, \quad \Delta p_z \, \Delta z \gtrsim \hbar.$$

Formula (4.3.2) is called the **uncertainty relation**. It is quite significant and applicable not only to electrons but also to other microparticles.

Let us consider certain corollaries of the uncertainty relation. Assuming that the coordinate of an electron is known precisely, i.e. $\Delta x \to 0$, we find from (4.3.2) that $\Delta p_x \to \infty$, i.e. the momentum and hence the velocity v become indeterminate both in magnitude and in direction. In other words, *the coordinate and velocity of a particle cannot have strictly determinate values at the same instant of time.*

This circumstance means that if the coordinate of an electron is known at a certain instant of time t, the coordinate becomes quite indeterminate at the instant $t + \Delta t$. Indeed, $v \, \Delta t$ is the displacement of the electron over time Δt, and since the magnitude and direction of velocity v are indeterminate at the given instant of time, the electron coordinate also turns out to be indeterminate. We can speak only of the probability of an electron being located at a certain point in space. The concept of the electron trajectory[3] under these conditions becomes meaningless.

Other examples of the uncertainty relation are (i) between the uncertainties ΔE in energy and Δt in time:

$$\Delta E \, \Delta t \gtrsim \hbar, \tag{4.3.3}$$

(ii) between the uncertainty ΔL_z in the projection of the angular momentum on the Z-axis and $\Delta\varphi_z$ in the angle of rotation about the Z-axis:

$$\Delta L_z \, \Delta\varphi_z \gtrsim \hbar \tag{4.3.4}$$

and, similarly,

$$\Delta L_y \, \Delta\varphi_y \gtrsim \hbar, \quad \Delta L_x \, \Delta\varphi_x \gtrsim \hbar.$$

4.4. Probability Waves

The waves associated with microparticles are sometimes called **probability waves**. These waves cannot be regarded as a spatial distribution of a field like the electromagnetic

[3] See Sec. 4.4.3 for the limits of applicability of the concept of trajectory.

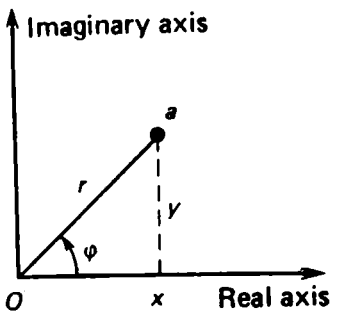


FIG. 4.6

or some other field. They cannot be measured directly and are described, as a rule, by complex (and not real) functions.

4.4.1. Complex Numbers. Euler's Formula. Apparently, the reader has a vague notion about complex numbers. We shall recall some of these ideas. A **complex number** a is a number of the type $a = p + iq$, where p and q are real numbers, and i is an imaginary unit defined as $i^2 = -1$. Each complex number a has a **complex conjugate number** $a^* = p - iq$ such that

$$aa^* = (p + iq)(p - iq) = p^2 + q^2. \tag{4.4.1}$$

Complex numbers can be represented graphically by points in a "complex plane" (Fig. 4.6). In this case, any complex number can be presented as

$$(x + iy) = r(\cos\varphi + i\sin\varphi)$$

since

$$x = r\cos\varphi, \quad y = r\sin\varphi. \tag{4.4.2}$$

Representation (4.4.2) of complex numbers is called their **trigonometric form**. This form is convenient for the multiplication of complex numbers:

$$a_1a_2 = r_1r_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)]. \tag{4.4.3}$$

In particular,

$$aa \equiv a^2 = r^2(\cos 2\varphi + i\sin 2\varphi). \tag{4.4.4}$$

Here $r = |a|$ is called the **modulus of a complex number** and is equal (see Fig. 4.6) to $\sqrt{x^2 + y^2}$. Consequently, we can write (4.4.1) in the form

$$aa^* = |a|^2. \tag{4.4.5}$$

Thus, *the square of the modulus of a complex number is a real number.*

Let us prove the possibility of the existence of the formula **(Euler's formula)**

$$e^{i\varphi} = \cos\varphi + i\sin\varphi, \tag{4.4.6}$$

where e is a real number defined as the limit of the expression

$$e = \left(1 + \frac{1}{n}\right)^n \tag{4.4.7}$$

as $n \to \infty$. To prove Euler's formula, we write (4.4.7) in the form

$$e^{1/n} = 1 + \frac{1}{n}. \tag{4.4.8}$$

This equation can be written for small $\varphi = 1/n$ in the form

$$e^{\varphi} \approx 1 + \varphi. \tag{4.4.9}$$

This relation can be generalized to the form

$$e^{i\varphi} \approx 1 + i\varphi. \tag{4.4.10}$$

Turning to (4.4.6), we note that for small φ, the function $\cos\varphi \simeq 1$ and $\sin\varphi \approx \varphi$, i.e. the right-hand side of (4.4.6) is also equal to $1 + i\varphi$. Thus we have verified the validity of formula (4.4.6) for small angles. It is also found to be true for all values of φ.

With the help of Euler's formula, we can present a complex number in exponential form. Indeed, multiplying (4.4.6) by r, we directly obtain this representation:

$$a = re^{i\varphi}, \tag{4.4.11}$$

where r is the modulus of the complex number a. In view of the fact that

$$e^{i\varphi} \cdot e^{-i\varphi} = e^0 = 1,$$

relation (4.4.5) is automatically satisfied.

4.4.2. Complex Probability Waves. The Superposition Principle. The probability waves describing the behaviour of an electron have the form

$$\Psi = \psi e^{i\varphi}. \tag{4.4.12}$$

In mathematical sense, probability waves as complex functions are characterized by two real functions, viz. the

modulus $|\psi|$, or the **amplitude of the probability wave**, and the **phase** φ. The phase reflects the wave properties of the electron (it "governs" the interference of the waves associated with the electron), while the amplitude reflects its corpuscular properties. To be more precise, the probability distribution for the electron to be situated at a certain point in space is given by the expression

$$\Psi\Psi^* = |\Psi|^2 = |\psi|^2. \tag{4.4.13}$$

Naturally, $|\psi|^2$ can only be positive, finite, single-valued, and continuous. The sum of the values of $|\psi(x, y, z)|^2$ for all possible points in space must be equal to unity since the electron must be situated somewhere.

Such a mathematical model has been constructed in quantum mechanics in the form of the **wave function** Ψ combining the corpuscular properties of particles and the wave nature of motion delectably. The probabilistic interpretation of the wave function turned out to be the only correct, logical, and noncontradictory interpretation.

With the help of probability waves, we can describe the experiment on the diffraction of electrons (see Fig. 4.4). When hole *1* is open (and hole *2* is closed), the electron is described by the function

$$\Psi_1 = \psi_1 e^{i\alpha_1}.$$

The probability of the electron to be located at different points on the screen is given by the expression

$$|\Psi_1|^2 = |\psi_1|^2.$$

If hole *2* is open (and hole *1* is closed), the electron is described by the function

$$\Psi_2 = \psi_2 e^{i\alpha_2}.$$

The probability of its being located at points on the screen is given by

$$|\Psi_2|^2 = |\psi_2|^2.$$

When both holes are open, the electron state is described by the function

$$\Psi_{12} = \Psi_1 + \Psi_2. \tag{4.4.14}$$

For the probability of electron coordinates, we obtain the relation

$$|\Psi_{12}|^2 = |\Psi_1 + \Psi_2|^2$$
$$= |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1|\,|\psi_2| \cos(\alpha_1 - \alpha_2). \qquad (4.4.15)$$

Comparing this relation with (3.9.3), we find that the probability distribution for electrons incident on the screen follows an interference pattern. We use Euler's formula (4.4.6) in writing Eq. (4.4.15), i.e. we take the real part of $e^{i(\alpha_1-\alpha_2)}$ which is expressed in the form $\cos(\alpha_1 - \alpha_2)$.

The connection between the electron states expressed through Eq. (4.4.14) is called the **superposition principle of quantum states**. This is one of the basic principles of quantum mechanics. For the corpuscular interpretation of this principle, let us write it in a more general form (as before, we consider the passage of an electron through a screen with two holes):

$$\Psi_{12} = a_1\Psi_1 + a_2\Psi_2,$$

where a_1 and a_2 are certain (generally complex) coefficients. If the system is in the state Ψ_1 (only hole *1* is open), then $a_1 \neq 0$ and $a_2 = 0$; here, $|a_1|^2 = 1$, i.e. the probability of the electron passing through hole *1* becomes a certainty. If, however, the system is in the state Ψ_2 (only hole *2* is open), then $a_2 \neq 0$, $a_1 = 0$; naturally $|a_2|^2 = 1$. When the system is in the state Ψ_{12} (both holes are open), the electron can pass through hole *1* with probability $|a_1|^2$ or through hole *2* with probability $|a_2|^2$. Moreover,

$$|a_1|^2 + |a_2|^2 = 1.$$

Note that in the idealized experiment on diffraction of electrons, the distance between holes *1* and *2*, as well as the diameters of the holes, are comparable with the wavelength λ of the probability wave. The probability wave associated with the electron as if passes through both holes and interferes with itself, ensuring a statistical appearance of electrons at different places. It is in this sense that we speak of each electron as interfering with itself.

4.4.3. Limiting Transition to Classical Mechanics. The interrelation between quantum and classical mechanics is

analogous in many respects to the interrelation between wave and geometrical optics. In quantum mechanics, the dynamic equation of motion describes the state of a particle in space and time through a wave function. This equation is called the **Schrödinger equation**. This is a linear differential equation, which is in keeping with the superposition principle. The motion of a particle in classical mechanics is treated as a motion along a trajectory.

In electrodynamics, the wave optical phenomena are described with the help of linear differential equations (Maxwell's equations). In the limiting case of geometrical optics, light is assumed to propagate along certain trajectories, i.e. rays.

Under normal conditions, the wavelength λ of light waves is much smaller than the size d of bodies with which light interacts. It was shown in Sec. 3.9.3 that the wave properties of light (interference and diffraction) cannot be observed in this case.

One can naturally expect that the concept of trajectory of a particle can be applied in quantum mechanics also if the wavelength λ of the probability wave associated with a particle is much smaller than the size d of the characteristic region of motion. Let us verify this by considering a few examples. We shall first determine the wavelength λ of the wave associated with an electron moving in a cathode-ray tube under a voltage of 100 V. We use expression (4.3.1) $\lambda = 2\pi\hbar/p$, in which the electron momentum p is determined from its relation with the energy $E = p^2/(2m_e)$. This gives

$$\lambda = \frac{2\pi\hbar}{\sqrt{2m_eE}} .$$

Substituting the numerical values, we get

$$\lambda = \frac{6.3 \times 1.05 \times 10^{-34}}{(2 \times 10^{-30} \times 100 \times 1.6 \times 10^{-19})^{1/2}} \text{ m} \simeq 1.1 \times 10^{-10} \text{ m}.$$

This shows that the coordinates of an electron moving in a vacuum tube of diameter, say, $d = 0.1$ m are determined to within 10^{-10} m, which is an extremely high accuracy.

Unlike the case $\lambda \ll d$, when the concept of trajectory of a particle is justified and acceptable, the particle moves as a wave for $\lambda \sim d$. For example, it will be shown in Sec. 4.5.1 that an electron in an atom has a coordinate uncertainty

of the order of 10^{-10} m, which is comparable with the size of the atom itself. Obviously, the concept of trajectory becomes meaningless for an electron in an atom.

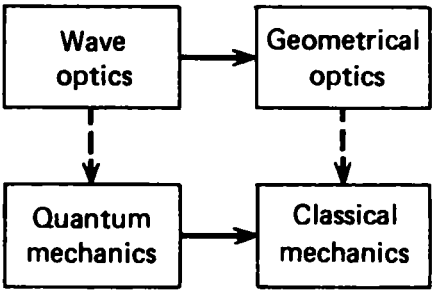


FIG. 4.7

When charged particles move in a Wilson's cloud chamber or in a bubble chamber, their trajectories can clearly be seen in the form of tracks. The accuracy with which the coordinates of a particle can be determined in this case depends on the size of the liquid or gas bubbles forming the track. The spatial resolution of tracks is of the order of 10^{-4} m, while the wavelengths associated, say, with relativistic electrons are $\lambda \sim \hbar/(mc) \sim 10^{-13}$ m.

If the expression describing the behaviour of a system contains the constant $\hbar$, it clearly indicates that the system is quantum mechanical. The relative value of $\hbar$ in comparison with other quantities having the same dimensions (Et or pr) determines the "degree of quantization" of a system. If the constant $\hbar$ can be eliminated, a transition from the quantum to the classical case is carried out automatically.

Thus, quantum mechanics contains classical mechanics as a limiting case. The limiting form of the probability wave for a "nearly classical" system contains the ratio $S/\hbar$ as phase, where S is called the **action** of the system, and $\hbar$ is treated as the **quantum of action**. For quasi-classical systems, the ratio $S/\hbar$ is quite large. In accordance with the optical analogy (Fig. 4.7) and formula (3.9.4), the wavelength λ of the probability wave in this case is small in comparison with the size d of the characteristic region of motion.

4.5. Electron in an Atom

4.5.1. Energy and Its Quantization. Let us consider the simplest atom, viz. the hydrogen atom. We proceed from the usual classical expression for the total energy

$$E = \frac{p^2}{2m} - \frac{e^2}{r}, \qquad (4.5.1)$$

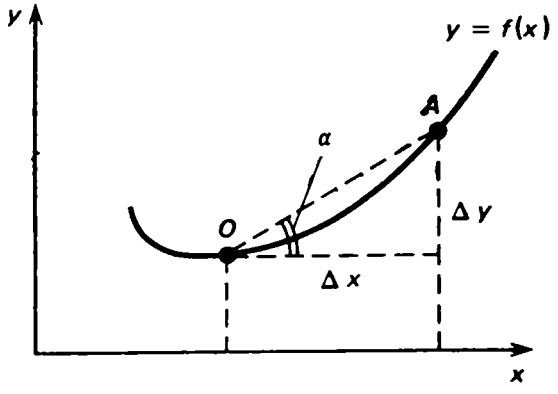


FIG. 4.8

where e is the electron charge, $p^2/(2m)$ is the kinetic energy of the electron, and the term $-e^2/r$ stands for the potential energy of interaction of the electron with the proton (cf. Eq. (3.4.10)).

In the quantum case, the coordinate and the momentum cannot exist simultaneously, hence we must use the uncertainty relations (4.3.2) in expression (4.5.1). For an atom, $\Delta p \sim p$ and $\Delta r \sim r$. Hence, substituting $p \sim \hbar/r$ into (4.5.1), we obtain

$$E = \frac{\hbar^2}{2mr^2} - \frac{e^2}{r}. \tag{4.5.2}$$

In order to proceed further, we shall require some mathematical concepts. The concept of **derivative** is introduced in high school at present, and the students are taught the **differentiation** of elementary functions. For readers who are not acquainted with these "subtleties", we shall provide a basic description of these concepts. To understand the meaning of the term "derivative", we turn to the plot of a function $y = f(x)$ (Fig. 4.8). Suppose that point O is fixed, while point A slides towards point O. In this case,

$$\frac{\Delta y}{\Delta x} = \tan \alpha.$$

In the limit, we obtain

$$\lim_{\Delta x \to 0} \left(\frac{\Delta y}{\Delta x}\right) \equiv \frac{dy}{dx} = \text{slope at point } O. \tag{4.5.3}$$

The quantity dy/dx is called the derivative of the function $y = f(x)$. Its calculation follows from the very definition:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{4.5.4}$$

By way of an example, let us determine the derivative of the function $y = x^2$. For this purpose, we can write

$$\frac{d(x^2)}{dx} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{x^2 + (\Delta x)^2 + 2x\,\Delta x - x^2}{\Delta x} = \lim_{\Delta x \to 0} (2x + \Delta x) = 2x.$$

The generalization of the formula

$$\frac{d}{dx}(x^2) = 2x$$

to the case of an arbitrary exponent n (integer, fraction, or negative) is the expression

$$\frac{d}{dx}(x^n) = nx^{n-1}. \tag{4.5.5}$$

For the particular case $n = 2$, we obtain the above expression from this relation.

Sometimes, it is required to find the minimum of a function $y = f(x)$. Proceeding from the geometrical meaning of the derivative (see Eq. (4.5.3)), we can conclude that the minimization condition for the function is

$$\frac{dy}{dx} = 0. \tag{4.5.6}$$

In the context, let us consider Fig. 4.9. The reader may observe at this stage that condition (4.5.6) is also the maximization condition for the function. This is indeed so! When a physical problem is considered, the formulation of the

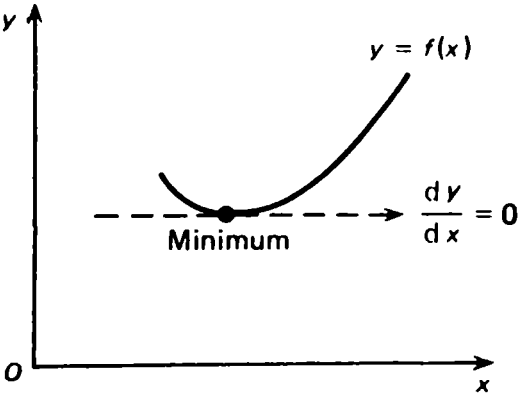


FIG. 4.9

problem itself indicates whether we are dealing with a maximum or a minimum.

Let us now return to the case of the electron in the hydrogen atom. Equation (4.5.2) does not provide any idea about the size of the atom. However, any physical system in free state must have the minimum possible energy. Applying condition (4.5.6) to the function $E(r)$ and taking the derivative of (4.5.2), we obtain

$$\frac{dE}{dr} = -\frac{\hbar^2}{mr^3} + \frac{e^2}{r^2} = 0.$$

While determining the derivative dE/dr, we have used formula (4.5.5). Solving the equation $dE/dr = 0$, we can determine the size of the atom (the **Bohr radius**)

$$r_a = \frac{\hbar^2}{me^2}. \tag{4.5.7}$$

This equation contains only fundamental constants whose combination establishes the real scale of lengths for atomic phenomena. Since Eq. (4.5.7) is written in CGS units, the constants must also be taken in CGS units, and a transition to SI units should be made only in the answer. As a result, we get

$$r_a = 0.52 \times 10^{-10} \text{ m}.$$

Substituting (4.5.7) into (4.5.2), we obtain the energy of an atomic electron:

$$E_a = -\frac{me^4}{2\hbar^2} = -13.6 \text{ eV}. \tag{4.5.8}$$

Thus, nature has prepared a true energy scale for atomic phenomena. The minus sign in this equation indicates that the electron is bound in the atom (see also Fig. 3.29). In order to detach the electron (to ionize the hydrogen atom), an energy of 13.6 eV is required.

Electrons in atoms are nonrelativistic objects. Indeed, from the uncertainty relation, the electron velocity $v \sim \hbar/(mr_a) \sim 10^6$ m/s, which is smaller than the ultimate velocity c by two orders of magnitude. Hence Eqs. (4.5.7) and (4.5.8) do not contain the fundamental constant c.

As far as the wavelengths λ of the probability waves are concerned, we obtain for atomic electrons $\lambda \sim \hbar/(mv) \sim 10^{-10}$ m, i.e. the uncertainty in their coordinate ($\Delta r \sim \lambda$) is of the order of the atomic size. The electron can be detected at any point in the atom, but its most probable location is close to the value r_a.

The application of the basic quantum-mechanical concepts on the wave-particle duality to an atomic electron naturally leads to a discretization (quantization) of the values of quantities like energy and angular momentum.

Stationary states of an electron in an atom, i.e. states with a definite energy, are described with the help of probability amplitudes $\psi(x, y, z)$ which are obtained from the time-independent Schrödinger equation.

For the stationary states of an atomic electron, we cannot even imagine the electron to move around the nucleus in the usual sense. The probability density $|\psi|^2$ varies from point to point. This means that the probabilities of detecting an electron in different volume elements in the space of an atom are different.

Since the spatial region for electron probability waves is somehow restricted to the atomic size, it can be expected that ψ form a certain discrete set of functions $\psi_n(x, y, z)$, where n runs through the values 1, 2, 3, ... Here, it is quite appropriate to draw an analogy with standing electromagnetic waves formed in a closed cavity (see Sec. 3.8.4 and Fig. 3.40).

The set of functions ψ_n whose explicit form will be given below has a discrete set of the corresponding energy values

$$E_n = \frac{E_a}{n^2} = -\left(\frac{me^4}{2\hbar^2}\right)\frac{1}{n^2}, \qquad (4.5.9)$$

where $n = 1, 2, 3, \ldots$ is called the **principal quantum number.**

Formula (4.5.9) represents the spectrum of quantized values of the electron energy in a one-electron atom. The number of "energy levels" in this spectrum is unlimited. For $n = 1$, this formula is transformed into Eq. (4.5.8). This value of energy corresponds to the **ground state** of the atom. In the ground state, the atom has the lowest energy. All states of the atom for $n > 1$ are called **excited states** (Fig. 4.10).

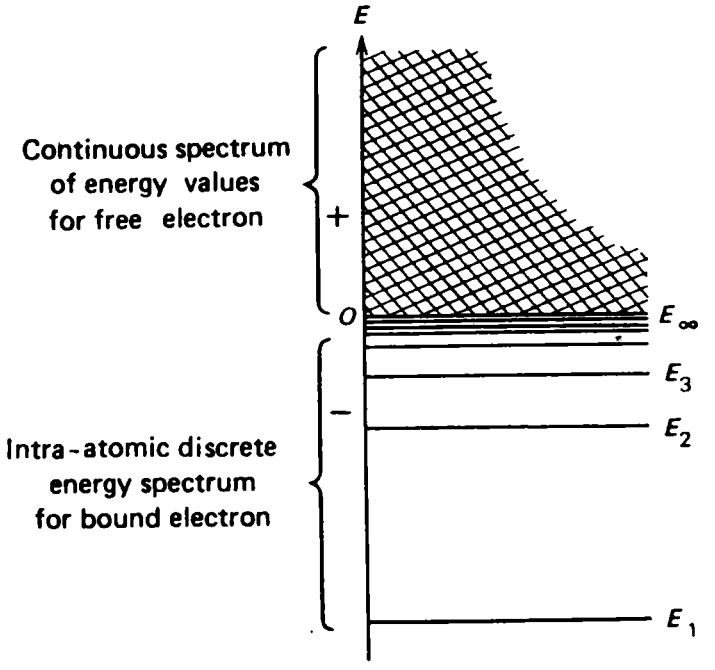


FIG. 4.10

4.5.2. Angular Momentum and Its Quantization. While considering the classical motion of a charge in a Coulomb field in Sec. 3.6.3, we obtained the following expression (in polar coordinates) for the total energy:

$$E = \frac{m}{2} v_r^2 + \frac{L^2}{2mr^2} - \frac{e^2}{r}, \tag{4.5.10}$$

where $mv_r^2/2$ is the radial part of the kinetic energy, and $L^2/(2mr^2)$ is its orbital part. Applying the uncertainty relation to Eq. (4.5.10) and considering that $mv_r r \sim \hbar$, we obtain

$$E = \frac{1}{2mr^2}(\hbar^2 + L^2) - \frac{e^2}{r}. \tag{4.5.11}$$

It can be seen from this equation that when E is quantized, the right-hand side can assume discrete values corresponding to E only on account of L (angular momentum). It can also be seen from (4.5.11) that $\hbar$ is a natural unit for measuring the angular momentum of the electron. It is found that

$$L^2 = \hbar^2 l\,(l + 1), \tag{4.5.12}$$

where l is called the **orbital quantum number** and has discrete values

$$l = 0, 1, 2, \ldots, (n-1).$$

Hence *the angular momentum of an electron in an atom is also quantized.*

A remarkable feature of the atomic electron is that it can have $L = 0$. From the viewpoint of classical mechanics, this corresponds to the motion of the electron along a radius passing through the centre (see Eq. (4.5.10)). The possible values of l are limited by the value of n.

The angular momentum $\mathbf{L}$ is a vector. This means that it is specified not only by its magnitude L but also by its direction. Naturally, the quantization of the values of L also entails a discretization of the direction of the angular momentum.

It should be recalled that in classical electrodynamics the mechanical angular momentum $\mathbf{L}$ is connected to its magnetic moment $\mathbf{p}_m$ (see Eq. (3.5.6)). In the quantum-mechanical case of an atomic electron, the orbital angular momentum and the corresponding magnetic moment are connected through the same relation

$$\mathbf{p}_m = \frac{e}{2mc}\mathbf{L}. \tag{4.5.13}$$

Here, we have considered that the atomic electron has a velocity $v \ll c$. In an external magnetic field $\mathbf{H}$, the magnetic moment $\mathbf{p}_m$ is oriented. The additional energy of such an electron in a magnetic field is determined with the help of a formula analogous to (3.4.18) for the energy of a dipole $\mathbf{d}$ in an electric field $\mathbf{E}$, i.e.

$$U_m = \mathbf{p}_m \cdot \mathbf{H} \sim \frac{e\hbar}{2mc} H \sim 10^{-3} \text{ eV}. \tag{4.5.14}$$

In this equation, we have taken the maximum possible value of H, viz. $\sim 10^5$ G, that can be obtained in laboratories. Keeping in view Eq. (4.5.8), we can conclude that the possible quantum states of an electron with a given n and different values of l form a group of closely spaced substates.

Let us now consider the discretization in the orientation of the angular momentum. Suppose that we select the Z-di-

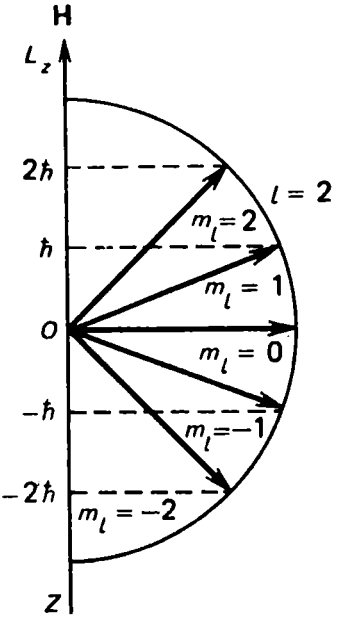


FIG. 4.11

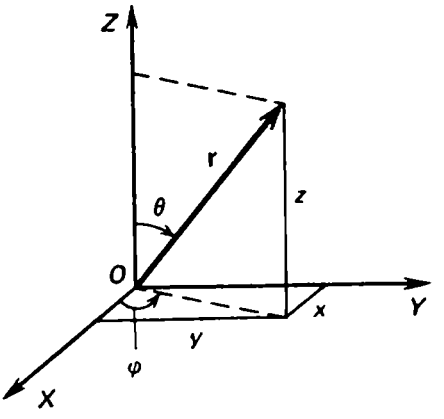


FIG. 4.12

rection (direction of the magnetic field). The projections of the angular momentum are then given by

$$L_z = m_l \hbar. \tag{4.5.15}$$

Only such orientations of the angular momentum are possible whose projections on the chosen direction in space are integers (positive and negative) in units of $\hbar$ (including zero). The number m_l is called the **magnetic quantum number**. Its values for a given value of l vary from $+l$ to $-l$ passing through zero. Hence the number of possible orientations of the angular momentum vector **L** in a magnetic field is $2l + 1$. As an example, let us consider the state of an electron with $l = 2$. The total number of possible orientations of the angular momentum is $2 \times 2 + 1 = 5$. The rest is shown in Fig. 4.11.

Note that we are talking only about one projection of the angular momentum since the other two are indeterminate. This follows from the uncertainty relation (4.3.4).

4.5.3. Probability Amplitudes and Quantum Numbers. A consistent analysis of the problem on the hydrogen atom involves the solution of the dynamic equation of quantum mechanics (the Schrödinger equation). The solution of this

equation is a set of probability amplitudes $\psi_{n,\,l,\,m}(x,\,y,\,z)$ which provide a complete description of the quantum states of the electron in the atom.

As usual, mathematical analysis can considerably be simplified by taking into consideration the symmetry conditions of the problem. The field of the nucleus is spherically symmetric. We make a transition from the Cartesian coordinates $(x,\,y,\,z)$ to **spherical coordinates** $(r,\,\theta,\,\varphi)$ as shown in Fig. 4.12. These coordinates are most suitable for describing physical situations involving spherical symmetry. In spherical coordinates, the probability amplitudes $\psi\,(x,\,y,\,z)$ are transformed into $\psi\,(r,\,\theta,\,\varphi)$. The situation is simplified in view of the fact that the function of three variables $\psi\,(r,\,\theta,\,\varphi)$ can be split into three functions $R\,(r)$, $\Theta\,(\theta)$, and $\Phi\,(\varphi)$, each of which depends only on one variable:

$$\psi\,(r,\,\theta,\,\varphi) = R\,(r)\,\Theta\,(\theta)\,\Phi\,(\varphi). \tag{4.5.16}$$

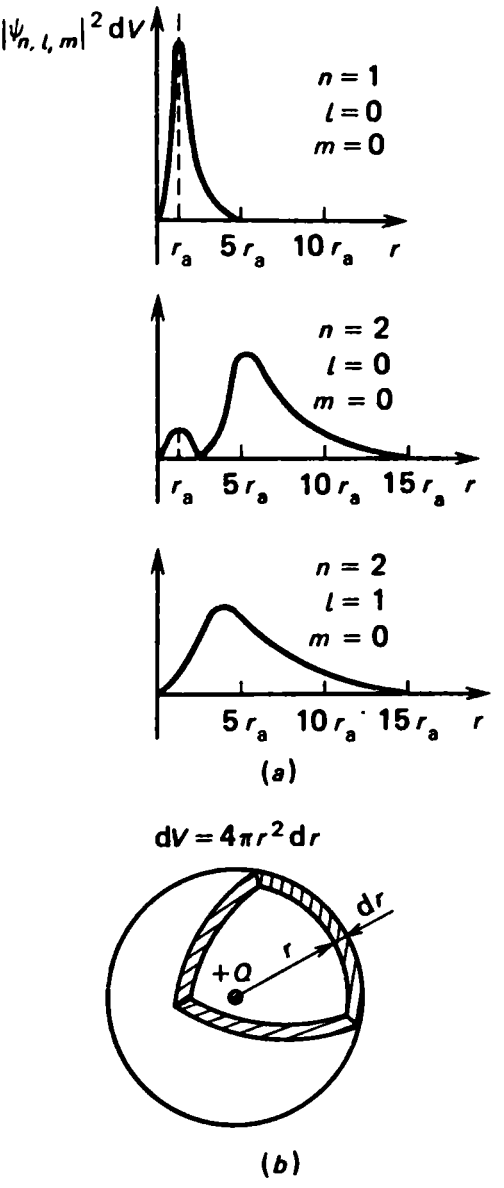


FIG. 4.13

The solutions of the Schrödinger equation must satisfy the following natural conditions:

(i) the probability amplitude at each point in space must be characterized by a single value;

(ii) the probability amplitude must not turn to infinity at any place;

(iii) the probability amplitude must be a continuous function of coordinates.

Such solutions are possible in the hydrogen atom problem only for integral values of certain parameters on which the functions $R(r)$, $\Theta(\theta)$, and $\Phi(\varphi)$ depend. Remarkably, these parameters turned out to be the quantum numbers n, l, and m_l. This means that (4.5.16) can be written in the form

$$\psi_{n,\,l,m} = R_{n,\,l}\Theta_{l,\,m}\Phi_m. \qquad (4.5.17)$$

The probabilities $|\psi_{n,\,l,\,m}|^2\,dV$ of an electron appearing in a spherical layer enclosed between the values r and $r + dr$ for definite values of n, l, and m_l are shown in Fig. 4.13*a*. The method of defining the volume element dV is clear from Fig. 4.13*b* (it should be noted that $4\pi r^2$ is the surface of the sphere, and dr is the thickness of the spherical shell).

In the three-dimensional problem on an atomic electron, the electron has three degrees of freedom, and hence three quantum numbers are needed to describe its behaviour.

4.6. Many-Electron Atom

4.6.1. Spin of an Electron. It has been shown experimentally that in addition to the orbital angular momentum associated with its motion in space, an electron also has a certain intrinsic angular momentum which exists even in the reference frame in which the electron is at rest. This intrinsic angular momentum must be attributed to the manifestation of internal degrees of freedom of the electron since the three coordinates completely account for its spatial degrees of freedom. The intrinsic angular momentum of a particle is called its **spin**. Naturally, the spin has a certain magnetic moment corresponding to it.

Within the framework of the concepts described above, the spin of an electron must be introduced as an additional characteristic. The relativistic generalization of quantum mechanics naturally requires that the electron spin be taken into account. Since the electron velocity $v \ll c$ in an atom, spin must be weakly reflected in the energy spectrum of the atom. Indeed, the magnetic spin-orbit interaction, which is proportional to $1/c^2$ (see below), determines only the fine structure (splitting) of energy levels. However, the electron spin plays a decisive role in the problem of many-electron

atoms. Spin is the fourth quantum number describing the intra-atomic states.

The **spin (intrinsic) angular moment** S of an electron is described by the formula

$$S^2 = \hbar^2 s\,(s+1), \tag{4.6.1}$$

which is analogous to (4.5.12). In (4.6.1), the spin quantum number $s = 1/2$. The discreteness of the spin orientation is described by the expression

$$S_z = m_s \hbar. \tag{4.6.2}$$

This relation is analogous to (4.5.15). The number of spin orientations is $2s + 1 = 2(1/2) + 1 = 2$. The projections of spin on a chosen direction in space are given by

$$S_z = \pm \frac{\hbar}{2}. \tag{4.6.3}$$

The relation between the spin **S** and its magnetic moment $\mathbf{p}_{ms}$ has the form

$$\mathbf{p}_{ms} = \frac{e}{mc}\mathbf{S}. \tag{4.6.4}$$

A comparison of this equation with (4.5.13) shows that the ratio of the magnetic moment to the angular momentum is twice as large in the case of spin. This difference clearly follows from relativistic quantum equations for an electron in an external field.

Let us evaluate the fine structure of energy levels in a hydrogen atom. This fine structure is associated with the interaction of spin magnetic moment $\mathbf{p}_{ms}$ and orbital magnetic moment $\mathbf{p}_{ml}$ of the electron. We proceed from the classical expression for the energy of spin-orbit magnetic interaction:

$$U_{sl} \sim \frac{\mathbf{p}_{ms}\mathbf{p}_{ml}}{r_a^3}. \tag{4.6.5}$$

This formula has the same structure as (3.4.19). We make the substitutions

$$p_{ml} = \frac{e}{2mc} L_l = \frac{e\hbar}{2mc},$$

$$p_{ms} = \frac{e}{mc} L_s = \frac{e\hbar}{mc}$$

into (4.6.5) and also use expression (4.5.7) for the atomic radius r_a. As a result, we obtain

$$U_{sl} \sim \frac{me^8}{2\hbar^4}\left(\frac{1}{c^2}\right) = \left(\frac{e^2}{\hbar c}\right)^2 \frac{me^4}{2\hbar^2}. \qquad (4.6.6)$$

A comparison of (4.6.6) with (4.5.8) shows that these expressions differ only in that the former contains a dimensionless multiplier $\alpha = e^2/(\hbar c) \simeq 1/137$, called the **fine-structure constant**. Since $\alpha^2 \sim (10^{-2})^2 \sim 10^{-4}$, we obtain

$U_{sl} \sim 10^{-4} E_a$.

Note that there is a significant difference between the orbital angular momentum of an electron and its spin. In accordance with (4.5.12), $L \gg \hbar$ for large values of l, which indicates a transition to the concepts of classical mechanics. The magnitude of the electron spin is given, according to (4.6.1), by

$$S = \frac{\hbar}{2}\sqrt{3}.$$

As $\hbar \to 0$, the spin $S \to 0$, which means that *there is no classical limit for the spin angular momentum*. Hence, strictly speaking, the electron spin cannot be assigned a visual classical interpretation, for example, attributing it to the rotation of the electron about its own axis.

Spin is a purely quantum property of a particle. In addition to the mass and electric charge, spin is an important constant describing a particle. Most elementary particles, including leptons and a large number of baryons, have a spin equal to 1/2; the spin of π- and K-mesons is zero, while photons have a spin equal to unity. As far as the atomic nuclei are concerned, they have an integral or half-integral spin depending on whether the total number of nucleons constituting a nucleus is even or odd (see Sec. 1.1).

All that has been stated about spin is described in Fig. 4.14 (cf. Fig. 4.11).

4.6.2. Systems of Identical Particles. Quantum Statistics. So far, we have considered the probability amplitude of particles to be a function of coordinates. The introduction of the concept of spin means that the probability amplitude must be a function of coordinates as well as the spin variable

indicating the value of the spin projection on a chosen direction in space and running through a limited number of discrete values.

In a system of identical particles (say, electrons), the probability amplitudes have a special symmetry in view of the indistinguishability of such particles. Indeed, since the particles do not have a trajectory, we cannot track the movement of each of them. Hence *the states of the system obtained as a result of a simple transposition of particles must be physically identical.*

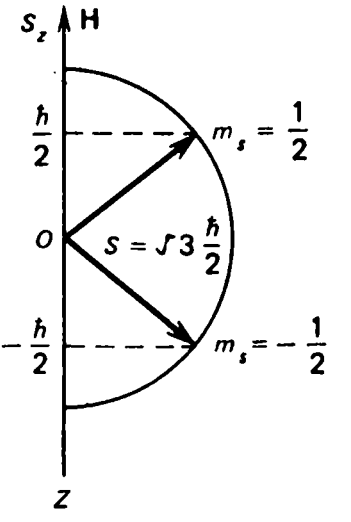


FIG. 4.14

Let us consider a system consisting of two identical particles with probability amplitude $\psi(\xi_1, \xi_2)$, where ξ_1 and ξ_2 denote the set of coordinates and spin of each individual particle. In particular, this notation indicates that the probability waves describing the behaviour of two or more particles are waves in an abstract multidimensional space rather than in the real three-dimensional physical space. This confirms once again that the probability waves cannot be associated with any real fields.

Interchanging the particles in the system under consideration, we obtain

$$|\psi(\xi_1, \xi_2)|^2 = |\psi(\xi_2, \xi_1)|^2 \tag{4.6.7}$$

since only the quantity $|\psi|^2$ has a real physical meaning. Consequently, it follows from (4.6.7) that

$$\psi(\xi_1, \xi_2) = \pm\psi(\xi_2, \xi_1). \tag{4.6.8}$$

This relation shows that a system formed by two identical particles can be described by a symmetric (+) or antisymmetric (—) probability amplitude relative to the transposition of the particles.

The above-mentioned property of symmetry of probability amplitudes is unambiguously connected to the spin of particles. Particles having a half-integral spin (in units of $\hbar$) are associated with antisymmetric ψ functions, while those

with zero or integral spin have symmetric ψ functions corresponding to them.

Before considering the consequences of the above-mentioned properties of transposition symmetry of probability amplitudes of systems consisting of identical particles, let us briefly touch upon a result from the probability theory: *the probability of simultaneous occurrence of two independent events is the product of the probabilities of individual events.* For example, let us cast dice, i.e. small cubes of bone marked with from one to six spots on each face. If one die is cast, the probability of each number appearing on top is 1/6. This result is quite obvious. If, however, two players independently cast a die each, the probability of simultaneous appearance of the same preset number on the two top faces is $(1/6) \times (1/6) = 1/36$. This can be easily verified by carrying out such trial casts singly and in pairs.

Let us now turn to a system of two identical noninteracting particles. Each of these particles can be in one of its quantum states ψ_{N1}, ψ_{N2}, etc., where N_i is the number of a quantum state of a particle. What will be the form of the probability amplitude ψ of the system? Firstly, it must be expressed as the product of the probability amplitudes of each of the particles in view of their independence; secondly, in accordance with the superposition principle (see Eq. (4.4.14)), ψ must contain a linear combination of these products differing only in the transposition of the particles. If the particles have integral spins, ψ must not change sign upon a transposition, i.e. should be represented by a symmetric combination

$$\psi(\xi_1, \xi_2) \sim \psi_{N1}(\xi_1)\,\psi_{N2}(\xi_2) + \psi_{N1}(\xi_2)\,\psi_{N2}(\xi_1). \qquad (4.6.9)$$

If, however, the particles have half-integral spins, ψ changes sign upon a transposition. In other words, it is an antisymmetric combination

$$\psi(\xi_1, \xi_2) \sim \psi_{N1}(\xi_1)\,\psi_{N2}(\xi_2) - \psi_{N1}(\xi_2)\,\psi_{N2}(\xi_1). \qquad (4.6.10)$$

Let us analyze this expression. If two of the numbers of quantum states are identical, then $\psi(\xi_1, \xi_2) = 0$. This means that *a system consisting of two identical particles with a half-integral spin cannot have more than one particle in one quantum state.*

It follows from (4.6.9) that *no restrictions are imposed on the filling of the quantum states by identical particles with an integral spin in a system.* The population rules of quantum states for particles with half-integral spins constitute **Fermi-Dirac statistics**, while those for particles with integral spins are termed **Bose-Einstein statistics.**[4]

These results can be directly generalized to systems containing any number of identical particles. The probability amplitudes of such systems can be symmetric or antisymmetric with respect to a transposition of any pair of particles. Thus, *in quantum mechanics, there is a sort of mutual effect of identical particles in a system even when there is no direct interaction force between the particles.*

4.6.3. Atomic Quantum States. Electrons in an atom form a system of identical particles with a half-integral spin. The aggregate of electrons in an atom obeys Fermi statistics, i.e. *each quantum state in the atom can be occupied only by one electron.* The quantum state of the atom itself is determined by the set of quantum numbers n, l, m_l, and m_s.

Let us determine the number of allowed quantum states for a given value of the principal quantum number n. The orbital quantum number l runs through integral values from $l = 0$ to $l = n - 1$ (see Eq. (4.5.12)). For each l, there are $2l + 1$ possible values of the magnetic quantum number m_l (see the remarks following Eq. (4.5.15)). Finally, by taking into consideration the spin magnetic quantum number m_s, we double the total number of states. Consequently, for a given n, the total number of allowed states is

$$2\sum_{l=0}^{n-1}(2l+1).$$

Let us calculate this sum. Successively assigning the above-mentioned values to l, we obtain

$$2\sum_{l=0}^{n-1}(2l+1)=2\{1+3+5+\ldots+[2(n-1)+1]\}$$
$$=2[1+3+5+\ldots+(2n-1)].$$

[4] For the sake of brevity, we shall henceforth refer to these statistics as Fermi statistics and Bose statistics respectively.

The terms in the brackets (their total number is n) form an arithmetic progression whose sum S_n is given by the expression

$$S_n = \frac{(a_1 + a_n)}{2} n.$$

Here, a_1 and a_n are respectively the first and the last term in the progression containing n terms in all. Applying this relation for S_n to the case under consideration, we obtain

$$S_n = \left[\frac{1 + (2n-1)}{2}\right] n = n^2.$$

Finally, we arrive at the relation

$$2 \sum_{l=0}^{n-1} (2l + 1) = 2n^2. \qquad (4.6.11)$$

The table for the values of quantum states presented below is comprised with the help of formula (4.6.11).

Table 4.1

Principal quantum number n	Number of quantum states in the atom corresponding to n
1	2
2	8
3	18
4	32

The above classification of atomic quantum states was obtained for the case of a one-electron atom. In a many-electron atom, there is an interaction between electrons. However, the "independent particle approximation" can be applied to these particles with a fairly high degree of accuracy. This means that the electrons must successively fill the atomic quantum states in accordance with the above classification.

The electron states corresponding to a given value of the principal quantum number (see Table 4.1) have quite close values of energy (see the remarks following Eqs. (4.5.14) and (4.6.6)). Such states are said to form a **shell**. If the electrons completely occupy all possible states in a given

shell, it is said to be completely filled. In a **closed shell**, the total orbital angular momentum of all electrons vanishes, as well as their total spin. The effective charge distribution in closed shells is spherically symmetric. This means that *the atoms in which all the shells are closed do not have an electric dipole moment.*

If we consider the chemical elements in the increasing order of their nuclear charge, a certain periodicity is observed in the variation of their chemical properties. The **Periodic Table of Elements** contains periods with 2, 8, 18, and 32 elements, which is in complete accord with the data represented in Table 4.1. This explains the origin of chemical periodicity.

4.7. Quantization of Atomic Radiation

4.7.1. Quantum Transitions. Line Spectra. From the above analysis and estimates, it can be easily concluded that a transition from an excited state to the ground state of an atom is accompanied by the emission of electromagnetic radiation. Since the energy of electron states in the atom is quantized, the emission energy is also quantized, i.e.

$$\Delta E = \hbar\omega, \tag{4.7.1}$$

where ΔE is the energy difference between the excited and the ground state.

It should be noted that transitions are also possible between excited states. If E_i and E_k are the energies of the ith and kth excited states, the emission frequency is given by

$$\omega_{ik} = \frac{E_i - E_k}{\hbar} . \tag{4.7.2}$$

Figure 4.15 shows the spectral series of the radiation emitted by a hydrogen atom. Naturally, the emission spectrum itself consists of lines corresponding to transitions between discrete energy levels in the atom. It can be seen from the figure that the excited atom may serve as a source of electromagnetic radiation with a very wide range of frequencies. The observed discrete radiation emitted by the atoms covers infrared, visible, and ultraviolet regions of the spectrum, as well as the X-ray region (in the case of heavy elements).

It should be observed that atoms of different chemical elements have different emission spectra. This forms the

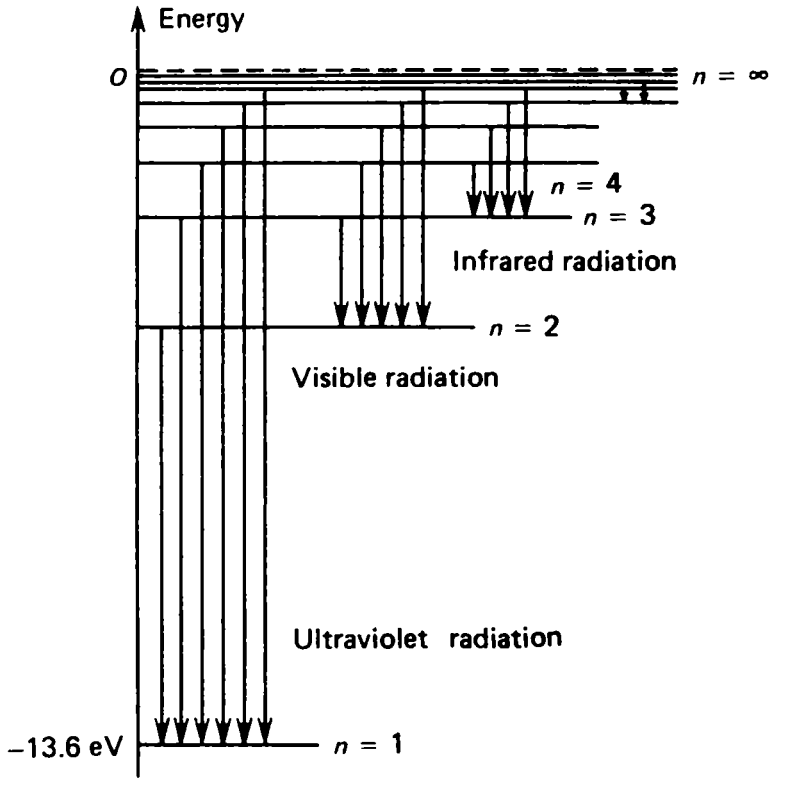


FIG. 4.15

basis of the spectral analysis of chemical composition of astronomical objects like the Sun and stars.

4.7.2. Photon. The Concept of Parity. Selection Rules. The energy of a radiation quantum emitted as a result of the transition

(Excited atom) → (Normal atom) + (Field quantum)

is given by

$$E = \hbar\omega, \tag{4.7.3}$$

while its angular momentum is given by

$$L = \hbar. \tag{4.7.4}$$

The electromagnetic radiation quantum, called a **photon**, is an elementary particle. However, it is different from ordinary elementary particles in several respects. For example, its total angular momentum cannot be divided into orbital and spin components. By definition, spin is the angular mo-

mentum of a particle at rest. A photon, however, cannot be at rest since it has a velocity c in all inertial reference frames.

Not all radiative transitions are possible in an atom, as it might appear from Eq. (4.7.2). Indeed, the energy and angular momentum conservation laws are obeyed for an isolated excited atom. A transition which is allowed from the energy point of view may turn out to be forbidden by the angular momentum quantization rules (4.5.12).

Experiments show that certain selection rules have to be applied to the set of energy and angular momentum quantum numbers in order to explain the radiative transitions in atoms. These selection rules follow from the angular momentum conservation law and the **parity conservation** for the quantum state of a closed system.

The parity conservation has a profound background. As a matter of fact, *physical relations do not change their form upon a coordinate transformation* $\mathbf{r} \to -\mathbf{r}$ (by considering Newton's laws of motion as an example, we shall verify this at a later stage). This transformation is called **coordinate inversion**. This circumstance means that *the laws of nature can be described by using both a left-handed and a right-handed coordinate system*[5] (Fig. 4.16*a*).

The transformation $\mathbf{r} \to -\mathbf{r}$ means, for example, that the velocity vector $\mathbf{v} = d\mathbf{r}/dt$ is transformed into vector $-\mathbf{v}$, the momentum vector $\mathbf{p} = m\mathbf{v}$ is transformed into $-\mathbf{p}$, while the force vector $\mathbf{F} = d\mathbf{p}/dt$ is transformed into $-\mathbf{F}$. Consequently, the equation of motion $m\, d\mathbf{v}/dt = \mathbf{F}$ in classical mechanics remains unchanged as a result of coordinate inversion.

The behaviour of physical quantities as a result of a coordinate transformation is described by using the concept of **parity**. We have seen that certain quantities like the vectors $\mathbf{v}$, $\mathbf{p}$, and $\mathbf{F}$ change sign upon a coordinate transformation $\mathbf{r} \to -\mathbf{r}$. Hence these quantities are assigned a **negative parity**. On the other hand, the angular momentum vector $\mathbf{L} = [\mathbf{rp}]$ defined as the vector product of $\mathbf{r}$ and $\mathbf{p}$ does not change sign upon a coordinate inversion. The moment of

---

[5] Figure 4.16*b* shows two operations: an inversion of the coordinate system and its rotation by 180° around the Y-axis. This is the process for obtaining a *mirror reflection*, i.e. for making a transition from a left-handed to a right-handed coordinate system.

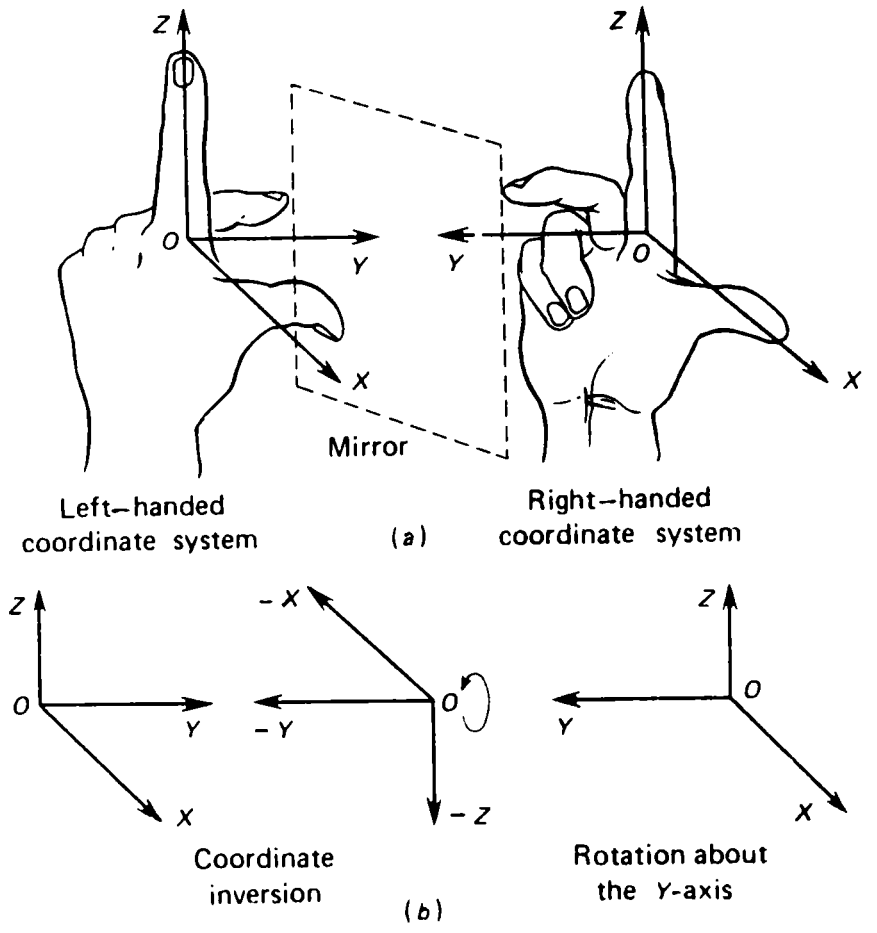


FIG. 4.16

force vector $\mathbf{K} = [\mathbf{rF}]$ also possesses the same property. Such vectors are said to possess a **positive parity**.

In quantum mechanics, the requirement of invariance of the equations to the inversion transformation imposes certain restrictions on the nature of the wave functions $\psi_{n,\ l,\ m}(x, y, z)$ describing the quantum states of the intra-atomic electrons. These functions can be even, i.e. $\psi(-x, -y, -z) = \psi(x, y, z)$, or odd, i.e. $\psi(-x, -y, -z) = -\psi(x, y, z)$ (Fig. 4.17).

What are the consequences of the mirror symmetry in the laws of physics? *If a certain process is allowed in nature, its mirror reflection is also allowed, and both processes have the same characteristics.* Processes occurring under the action of gravitational, electromagnetic, and strong interactions have a mirror symmetry (weak interactions are an exception to

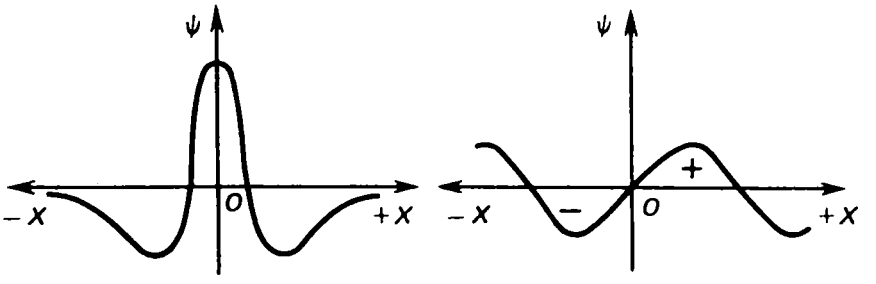


FIG. 4.17

this rule!). Such processes satisfy the selection rules associated with the parity conservation: *the parity of a closed system must not change as a result of the processes taking place in it.*

Let us consider the case of radiative transitions in an atom. By way of an example, we analyze the radiation emitted by a hydrogen atom upon a transition from an excited quantum state $2p$ ($n = 2$, $l = 1$) to the ground state $1s$ ($n = 1$, $l = 0$). The nature of symmetry and the spatial distribution of the electromagnetic field emitted during such a process are the same as of the field emitted by an oscillating electric dipole.

The photons corresponding to this type of radiation are called **electric-dipole photons**. As was mentioned above, the total angular momentum **j** of a photon cannot be divided into the orbital angular momentum **l** and the spin angular momentum **s**, and assumes only integral values $j = 1, 2, 3, \ldots$. For an electric-dipole photon, $j = 1$ and the parity P_γ of the state is equal to $(-1)^j = -1$, i.e. such a photon has an odd parity.

In the ground state, the hydrogen atom has an orbital quantum number $l = 0$ and the parity $P_{l=0}$ of this state is equal to $(-1)^l = (-1)^0 = +1$, i.e. it is a state with even parity. For the initial excited state $l = 1$, however, the parity $P_{l=1} = (-1)^l = -1$ is odd. Hence such an emission process can be represented in the form shown in Fig. 4.18.

As a result of such a radiative transition, the orbital quantum number of the atom changes by $\Delta l = -1$. Naturally, the angular momentum is conserved in this case.

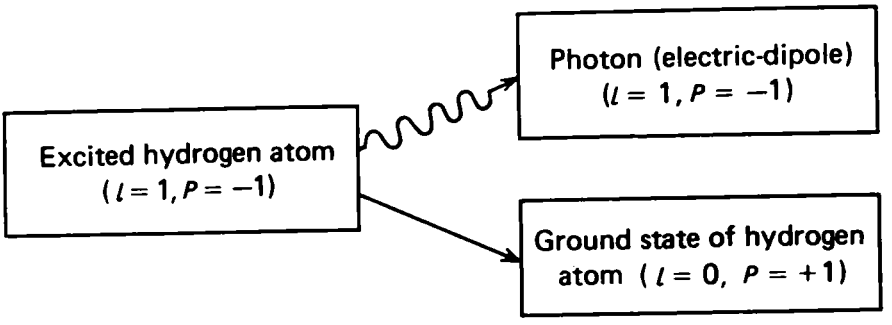


FIG. 4.18

So far as the parity conservation is concerned, it should be remarked that *the parity P of a system of two independent parts with parities* P_1 *and* P_2 *is equal to the product* P_1P_2 *of the parities of these parts.* In this case, we get

$$P_{l=1}^{\text{atom}} = P_\gamma P_{l=0}^{\text{atom}}.$$

Indeed, substituting the corresponding parity values, we obtain

$$(-1) = (-1)\,(+1).$$

Consequently, the parity conservation in *dipole* radiative transitions (in addition to the energy and angular momentum conservation) leads to the following **generalized selection rule:**

$$\Delta l = \pm 1, \tag{4.7.5}$$

i.e. *the angular momentum can change only by* $\hbar$.

For multipole radiative transitions, which have a much lower probability, this selection rule may not be observed.

Let us return to the properties of the photon. The fact that photons have an integral angular momentum means that a photon is a boson. This implies that *an aggregate of photons obeys Bose statistics.* This important circumstance has several remarkable consequences.

The emission of radiation by the atoms considered above may be called **spontaneous emission.** However, if an atom is exposed to a photon of such a type that it is capable of emitting, the probability of photon emission by the atom sharply increases. This follows from Eq. (4.6.9): the probability amplitude of forming a system of two identical

photons, i.e. photons having identical energy and angular momentum, is doubled. If there were N photons in the initial state of a field, the probability of emission of a similar photon by the atom increases N times. This effect is called **induced emission.**[6]

The applicability of Bose statistics to photons means that *an indefinite number of photons can be accommodated in the same quantum state.* On the one hand, this enables us to introduce the concept of field as an aggregate of photons, and on the other hand, it allows a limiting transition from the photon pattern of the field to the concept of classical electromagnetic waves.

These arguments directly lead us to an expression for the photon momentum **p**. Indeed, proceeding from relation (3.8.9) between the momentum density and the energy density of a classical electromagnetic wave, we obtain

$$p = \frac{E}{c} = \frac{\hbar\omega}{c} . \qquad (4.7.6)$$

The classical electromagnetic wave has a corresponding set of photons of the same frequency (energy) and polarization (angular momentum), which propagate in the same direction.

Thus, *quantum properties can be exhibited not only by ordinary particles but also by a field.* Photons are both quantum-mechanical and relativistic objects (see Eq. (4.7.6)). In this connection, it is appropriate to consider a schematic diagram of the interaction between different fundamental theories (Fig. 4.19). The arbitrary notation $\hbar = 0$ and $c = \infty$ in this figure indicates that the finite nature of these quantities is not significant for the phenomena under consideration.

Finally, let us consider another property of the photon as a relativistic particle. In accordance with the general relativistic relation (2.7.6), the expression $p = E/c$ means that the photon is a massless particle ($m = 0$).

In turn, this property of the photon leads to the conclusion that *photons can have a velocity only equal to c.* Indeed, if in formulas (2.7.9) and (2.7.10) we put $v = c$, the momentum

[6] An analysis of the applied problems associated with a schematic diagram of induced emission (see Sec. 7.7.9) permits a more detailed description of the concept of "identical" photon.

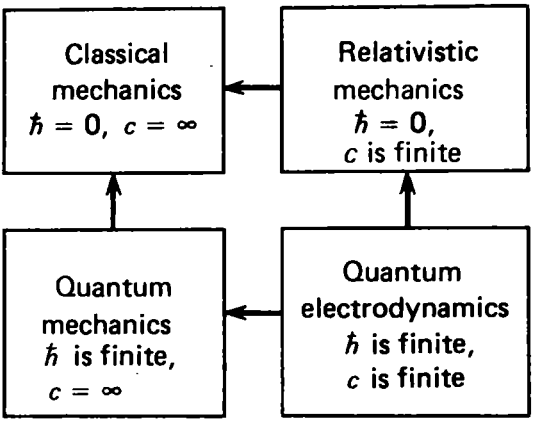


FIG. 4.19

p and energy E become infinite. This means that *a particle with a nonzero mass cannot move with a velocity equal to c.* However, if $m = 0$ and $v = c$, the energy and momentum relations (2.7.9) and (2.7.10) lead to indeterminate forms of the type 0/0 and may remain finite. The fact that this is indeed so is confirmed by quantum mechanics (see Eqs. (4.7.3) and (4.7.6)).

4.8. Photon-Electron Interaction. The Photoelectric Effect. The Compton Effect

Let us consider the **photoelectric effect** in a hydrogen atom. The whole process involves the ionization of the atom due to the absorption of a photon from an external radiation. Ultraviolet radiation having a wavelength $\lambda \sim 10^{-7}$ m was used in experiments on the photoelectric effect. In this case, the radiation frequency is

$$\omega = 2\pi \frac{c}{\lambda}, \quad \omega \simeq 6.3 \frac{3 \times 10^8}{10^{-7}} \text{ s}^{-1} \simeq 2 \times 10^{16} \text{ s}^{-1}.$$

The photon energy is given by

$$\hbar\omega \sim 10^{-34} \times 2 \times 10^{16} \text{ J} \simeq 2 \times 10^{-18} \text{ J}.$$

The ionization energy U_1 of a hydrogen atom is comparable with its ground state energy E_a, i.e.

$$U_1 \sim |E_a| = \frac{me^4}{2\hbar^2}.$$

This formula is written in the CGS units. Substituting the numerical values and going over to the SI units, we obtain $U_i \sim 10^{-18}$ J.

Consequently,

$$\hbar\omega \sim U_i \text{ (in the photoelectric effect).} \tag{4.8.1}$$

The electron velocity in an atom is $v \simeq 10^6$ m/s, i.e. $v \ll c$. Suppose that in the final state also, the free electron emitted by the atom has $v \ll c$. In other words, the problem is nonrelativistic as far as the electrons are concerned. In this case, the energy conservation law for the photoelectric effect can be written in the form

$$E_\gamma = U_i + E_e, \tag{4.8.2}$$

where $E_\gamma = \hbar\omega$ is the photon energy, U_i is the ionization energy of the atom, and $E_e = mv^2/2$ is the kinetic energy of the photoelectron. Equation (4.8.2) can be written in the form

$$\hbar\omega = U_i + \frac{mv^2}{2}. \tag{4.8.3}$$

This relation was proposed by Einstein.

In the photoelectric effect, the photon momentum is small in comparison with the momentum of the atomic electron. Let us verify this. The photon momentum is

$$p_\gamma = \frac{\hbar\omega}{c}, \quad p_\gamma \simeq \frac{10^{-34} \times 2 \times 10^{16}}{3 \times 10^8} \text{ kg·m·s}^{-1} \simeq 10^{-26} \text{ kg·m·s}^{-1},$$

while the atomic electron has a momentum

$$p_e = mv \sim 10^{-30} \times 10^6 \text{ kg·m·s}^{-1} \sim 10^{-24} \text{ kg·m·s}^{-1},$$

which means that

$$p_\gamma \ll p_e \text{ (in the photoelectric effect).} \tag{4.8.4}$$

Let us now consider the **Compton effect.** We shall be dealing with the scattering of a photon by a free electron. In this case, the experiment is carried out on very hard X-rays having a wavelength $\lambda \simeq 10^{-10}$ m corresponding to a frequency $\omega \simeq 2 \times 10^{19}$ s^{-1}.

Unlike the photoelectric effect, the scattering of photons by electrons is effective only when the photon momentum p_γ is comparable with the electron momentum p_e. Indeed,

$p_\gamma = \hbar\omega/c \sim 10^{-23}\ \text{kg}\cdot\text{m}\cdot\text{s}^{-1}$ and $p_e = mv \sim 10^{-24}\ \text{kg}\cdot\text{m}\cdot\text{s}^{-1}$, i.e.

$$p_\gamma \gtrsim p_e \text{ (in the Compton effect).} \tag{4.8.5}$$

As far as the energies of the particles are concerned, we are dealing with the relativistic case now, where the measure of energy is the rest energy (see Eq. (2.7.7)). In the present case, we are speaking of the rest energy of the electron:

$$mc^2 \sim 10^{-30}\,(3 \times 10^8)^2 \sim 10^{-13}\ \text{J}.$$

While the ionization energy $U_1 \sim 10^{-18}$ J for an atom is vanishingly small in comparison with the electron rest energy in the case of the photoelectric effect, i.e.

$$U_1 \langle\langle\langle\ mc^2 \text{ (in the photoelectric effect),} \tag{4.8.6}$$

the distinction between the energy of a quantum and the rest energy of an electron is considerably reduced in the Compton scattering:

$$\hbar\omega \ll mc^2 \quad \text{(in the Compton effect).} \tag{4.8.7}$$

Indeed, for X-ray quanta we have $\hbar\omega \sim 2 \times 10^{-15}$ J.

It should be noted that for $\hbar\omega \approx mc^2$ (to be more precise, for $\hbar\omega = 2mc^2$), a whole new range of relativistic quantum phenomena is observed: high-energy photons vanish, creating a particle-antiparticle pair in the process (in the present case, an electron-positron pair is formed; this phenomenon occurs in the vicinity of the atomic nucleus). The reverse process of annihilation of an electron and a positron accompanied by the emission of a photon is also possible.

Let us return to the Compton effect and calculate the frequency ω' of the scattered X-ray quantum. We proceed from the energy and momentum conservation laws during the scattering of a photon by a free electron at rest. Suppose that the phenomenon takes place in the plane of the figure (Fig. 4.20). In this case, we can write

$$\left.\begin{aligned} E_\gamma + E_{e0} &= E'_\gamma + E_e, \\ \mathbf{p}_\gamma &= \mathbf{p}'_\gamma + \mathbf{p}_e. \end{aligned}\right\} \tag{4.8.8}$$

Here, $E_\gamma = \hbar\omega$ is the energy of the incident photon with frequency ω, $E_{e0} = mc^2$ is the relativistic energy of the electron at rest (see Eq. (2.7.7)), $E'_\gamma = \hbar\omega'$ is the energy of the scattered photon with frequency ω', and E_e is the

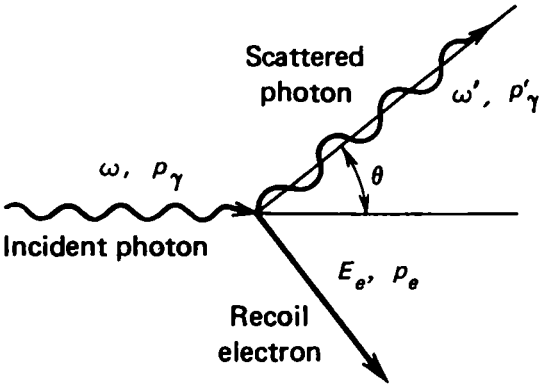


FIG. 4.20

total relativistic energy of the recoil electron (see Eq. (2.7.9)). Substituting these explicit expressions for the energies into the original equations (4.8.8), we obtain

$$\left.\begin{aligned} \hbar\omega + mc^2 - \hbar\omega' &= E_e, \\ \mathbf{p}_\gamma - \mathbf{p}'_\gamma &= \mathbf{p}_e. \end{aligned}\right\} \tag{4.8.9}$$

Next, we subject these equations to the following transformations. We divide both sides of the first equation by c and subtract the square of the second expression from the square of the first. This gives

$$\frac{1}{c^2}(\hbar\omega + mc^2 - \hbar\omega')^2 - (\mathbf{p}_\gamma - \mathbf{p}'_\gamma)^2$$

$$= \frac{E_e^2}{c^2} - p_e^2 \equiv m^2c^2. \tag{4.8.10}$$

The correctness of the right-hand side of this relation can be verified with the help of Eq. (2.7.6). In relation (4.8.10), $p_\gamma = \hbar\omega/c$ is the momentum of the incident photon, $p'_\gamma = \hbar\omega'/c$ is the momentum of the scattered photon, and $\mathbf{p}_e$ is the relativistic momentum of the recoil electron (see Eq. (2.7.10)). Recalling that the scalar product of two vectors is given by $(\mathbf{p}_\gamma \mathbf{p}'_\gamma) = p_\gamma p'_\gamma \cos\theta$ (see Fig. 4.20) and substituting the explicit expressions for p_γ and p'_γ into (4.8.10), we can present ω' in (4.8.10) in terms of the remaining quantities:

$$\omega' = \frac{\omega}{1 + \dfrac{\hbar\omega}{mc^2}(1 - \cos\theta)}. \tag{4.8.11}$$

This relation shows that the closer the quantity $\hbar\omega$ to the value mc^2, the more significant the decrease in the frequency of the scattered photon. This is the reason for using hard X-rays in the experiment.

From the viewpoint of quantum electrodynamics, the process of scattering of a photon by an electron can be interpreted as the absorption of a photon of energy $\hbar\omega$ by the electron followed by the emission of another photon with frequency ω' by the same electron, which thus acquires a momentum $\hbar(\omega - \omega')/c$.

Experiments on studying elementary processes, e.g. a direct measurement of the photon energy (in the photoelectric effect) and a direct measurement of the photon momentum (in the Compton effect) played a very significant role in the development of quantum-mechanical concepts and quantum electrodynamics.

4.9. Simultaneous Measurement of Quantities and the Concept of the Complete Set of Measurable Quantities

The analysis of this fundamentally important problem is a sort of generalization of the information gathered so far by us about the quantum-mechanical concepts.

Not all sets of physical quantities of different types can be measured simultaneously in quantum mechanics. This follows, among other things, from the uncertainty relations (4.3.2) according to which the momentum and the coordinate cannot have precise values at the same instant of time. This is due to the finite quantity, viz. Planck's constant $\hbar$, appearing in the uncertainty relations. The wave functions Ψ of the systems may have only the sets of simultaneously measurable quantities as parameters. It was shown that for electrons, such sets are formed by the following quantities:

$x, y, z, s_z,$

$p_x, p_y, p_z, s_z,$

$E_n, l, m_l, s_z.$

The set of all projections of the momentum (p_x, p_y, p_z) and the spin projection s_z on a certain direction Z is used for describing the states of a free electron. In this case, nat-

urally, the energy also has a definite value $E = (p_x^2 + p_y^2 + p_z^2)/(2m)$. The state of an electron bound in an atom is described with the help of a set of parameters including the energy E_n, the square of the magnitude of angular momentum L^2 (through the quantum number l), the projection L_z of the angular momentum (through the quantum number m_l), and the spin projection s_z.

The following sets are used for describing photons:

k_x, k_y, k_z, α,
E, L^2, L_z, P.

For freely propagating photons corresponding to plane classical waves, a set formed by three components of the wave vector $\mathbf{k}$ and the polarization α defining the orientation of the photon spin in space is used. The photon created in an atom as a result of quantum transitions is described by the set E, L^2, L_z, P, where P is the parity of the photon state.

For objects and systems studied by classical mechanics, the magnitude of action is much larger than the quantum of action ($S \gg \hbar$). Hence the restrictions imposed, say, by the uncertainty relations (4.3.2), on the simultaneous measurement of different types of physical quantities become insignificant. To be more precise, under conditions when the quantum of action can be neglected (i.e. when $\hbar \to 0$), the relations of the type (4.3.2) become invalid. Hence any set of physical quantities can be measured simultaneously, and the measuring process does not affect the object of measurement in any way.

4.10. Molecules

Atoms may combine to form a **molecule** if the energy of bound atoms is lower than the total energy of the isolated atoms. When a molecule is formed, the inner electron shells of the atoms remain practically unchanged, while the outer electrons are collectivized. Since the spatial region of motion of the collectivized electrons is broadened in this case, the uncertainty relations (4.3.2) indicate that their momentum decreases, i.e. the energy of the collectivized electrons also decreases.

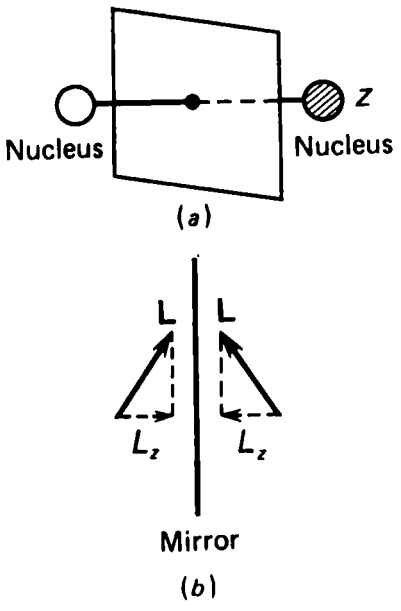


FIG. 4.21

Molecules of different types have different types of collectivization of their outer electrons (see Sec. 5.10.2).

On account of the large difference in the masses of electrons and nuclei in a molecule, the velocities of the nuclei are much smaller than those of the electrons. Hence, when considering the electron energy spectra in a molecule, we can assume that the nuclei are stationary. In the general case, the electron energy levels in a molecule are functions of the distance between the nuclei.

Diatomic molecules are of the simplest type. The total orbital angular momentum conservation law for all the electrons is not applicable in this case since the electric field of the two nuclei is not centrally symmetric. However, the electric field in this case does have an **axial symmetry** relative to an axis passing through both nuclei. A rotation of the entire system relative to this axis does not change the properties of the system. This means that the projection of the total angular momentum on this axis will be conserved, and we can classify the energy levels of molecules on the basis of this projection.

Besides axial symmetry, diatomic molecules also possess a symmetry relative to a plane passing through the axis (Fig. 4.21*a*). The energy state of a molecule remains unchanged as a result of mirror reflection (replacement of the coordinate **r** for all electrons by —**r**). However, if the angular momentum is nonzero, the sign of the projection of the angular momentum is reversed as a result of mirror reflection (Fig. 4.21*b*). Hence each energy value in the molecule has two quantum states corresponding to it. These states differ in the direction of projection of the angular momentum onto the axis.

Let us now consider the effect o f the nuclear motion on the energy spectrum of diatomic mol ecules. In such a molecule,

(i) the nuclei can perform small oscillations relative to the equilibrium position, and (ii) bound nuclei can rotate like dumb-bells about a certain point on the line joining these nuclei (naturally, in this case we must consider the rotation of the molecule as a whole). The vibrational and rotational motions in a molecule are naturally quantized. As a result, the electron energy levels E_{elec} of the molecule are split under the action of the vibrational energy E_{vib} of the nuclei, thus forming a **fine structure** of energy levels. In turn, these fine structure lines are also split owing to the rotational energy E_{rot} of the molecule as a whole, thus forming a **hyperfine structure** (bands). Thus, the energy of a diatomic molecule can be presented in the form

$$E = E_{\text{elec}} + E_{\text{vib}} + E_{\text{rot}}.$$

The term E_{elec} also includes the energy of electrostatic interaction between the nuclei forming the molecule.

The energy spectrum of the molecule is characterized by the following relation:

$$\Delta E_{\text{elec}} \gg \Delta E_{\text{vib}} \gg \Delta E_{\text{rot}}.$$

The energy difference ΔE_{elec} is of the order of several electron volts, $\Delta E_{\text{vib}} \sim 0.1$ eV, while $\Delta E_{\text{rot}} \sim 10^{-3}$ eV. The emission spectra of molecules resulting from transitions between electron states lie in the visible and ultraviolet regions. Quantum transitions between the vibrational levels are responsible for emission in the infrared region. Finally, quantum transitions between the rotational levels produce radiation in the microwave region (with wavelengths varying between 0.1 mm and 1 cm). An analysis of the emission spectra of molecules may provide detailed information about the molecules themselves.

Let us consider the interaction of atoms separated by large distances. We shall consider two identical atoms. Suppose that the outer electron shells of these atoms are closed and that the atoms are in the ground state. On account of the complete spherical symmetry of the states of such atoms, the mean electric dipole moment is zero. However, the instantaneous values of the electric dipole moments $\mathbf{d}$ of the atoms are nonzero. Consequently, the electrostatic bond (3.4.19) between the atoms leads to an interaction. This,

however, is an effect having second order of smallness, and hence

$$\begin{pmatrix}\text{Atom-atom}\\ \text{interaction energy}\end{pmatrix} \sim |U_{dd}|^2 \sim -\frac{\text{const}}{r^6}.$$

The notation used here is the same as in Eq. (3.4.20). This interaction is attractive in nature. The **(van der Waals) forces of interaction** act at much larger distances than the forces of chemical affinity. Unlike the latter type of forces, the van der Waals forces do not have the property of saturation and hence lead to the formation of the condensed state of matter (liquids and solids).

Chapter 5

MACROSCOPIC BODIES AS AGGREGATES OF PARTICLES. THERMAL PHENOMENA

5.1. The Basic Problem of Statistical Physics

Physical systems formed by an enormous number of particles (atoms or molecules) are called **macroscopic bodies**, or **macrobodies**. The very possibility of the existence of macroscopic bodies as spatially distributed stable aggregates of particles (nuclei and electrons) follows from quantum-mechanical concepts. Indeed, according to classical concepts, the electrostatic interaction between nuclei and electrons must result, even under normal conditions, in an overall condensation of particles into an overdense material. However, even in a system like an atom, an electron coming too close to the nucleus would have such a high momentum in view of the uncertainty relations that it could not be confined to the nucleus. The electron shells of neighbouring atoms cannot touch each other since electrons are Fermi particles.

Under normal conditions, macroscopic bodies contain about 10^{25}-10^{28} atoms/m^3. How can we describe the behaviour of such huge aggregates of particles? If we want to be consistent, we should try to apply the methods of (classical or quantum) mechanics. However, within the framework of, say, classical mechanics, such an aggregate of interacting particles can be described only if we (i) construct and solve a system of equations of motion containing as many equations as the number of interacting particles, i.e. of the order of, say, 10^{26}, and (ii) know the initial conditions (coordinates and velocities) of all the particles in the body. Obviously, such a problem cannot be solved even with the help of the most advanced computers.

The way out of this situation is indicated by mechanics itself. Indeed, let us recall the energy, momentum, and angular momentum conservation laws. For a closed system of particles, irrespective of the way in which the particles in-

teract and irrespective of the complexity of the manner in which they change their states, the total energy, momentum, and angular momentum of all the particles in a body must remain constant. If, for the sake of simplicity, we consider a macroscopic body in a reference frame in which it is at rest, we are left with only one invariable quantity, viz. the total internal energy of the macroscopic system.

Speaking of the energy of quantum systems, we usually mean that we are dealing with the energy spectra of atoms and molecules. However, quantum mechanics also uses the concept of the **energy spectrum of macroscopic bodies**. Only for ideal gases, i.e. for bodies consisting of noninteracting particles, the energy levels of a body are composed of the energy levels of the particles constituting it. For condensed bodies (liquids and crystals), the complexity and intensity of the interacting particles constituting such bodies warrant only a theoretical possibility of finding the overall energy spectrum of a body. Indeed, different collective forms of particle motion are realized in such bodies as a result of interaction between particles. In other words, the energy spectrum may contain a large number of branches. At the same time, many bodies are usually formed by subsystems of different particles, e.g. the ionic lattice and "free" electrons. The branches of energy spectra for each subsystem may influence one another. Such an influence may even cause a qualitative variation of the energy spectrum (an example of this is the transition in metals from normal to superconducting state). It is, however, remarkable that *for weakly excited states of a body, the energy levels may be represented in terms of the energy of the ground state and the sum of the energies of elementary excitations* (*quasi-particles*). We shall discuss these quasi-particles at a later stage.

Let us apply the principles of physical statistics[1] to the simplest bodies, viz. gases. Individual particles in such systems are subjected to the disordered action of the surrounding particles, and an isolated particle passes through all its possible states in the course of time. The only thing that one may try to determine for such a particle is the probability of its existence in each state.

Suppose that the gas under consideration remains a gas

---

[1] Also called statistical physics.

under all conditions. Then the existence of the body on the lowest macroscopic quantum level must correspond to a single microscopic distribution of particle states. When the body is excited (i.e. when it is transformed from the lowest to a higher energy level), the possibility of attaining this macroscopic state through different microscopic energy distributions of particles increases sharply. These distributions should be such that the total energy of the particles coincides with the total energy of the body. The main problem is that of finding the probability distribution for the mean number of particles occupying various quantum energy levels. In other words, *the main problem of statistical physics, which investigates the properties of macroscopic bodies as aggregates of particles, is to determine the "population density" of particles on different energy levels of the macroscopic system.*

5.2. Macroscopic Quantities. Fluctuations

Let us mentally divide the body under investigation into small (and yet macroscopic) subsystems. It is natural to characterize the state of the body in terms of the energy distribution over various subsystems. For a closed body, the sum of the energies of all the subsystems is unchanged in view of the energy conservation law. The total energy in the body can be distributed over different subsystems in different ways. Suppose that a certain subsystem is in a quantum state with energy E_n. Then, knowing the law of distribution of probabilities w_n of finding the subsystem in a certain state n, we can determine the mean value $\langle f \rangle$ of a quantity:

$$\langle f \rangle = \sum_n f_n w_n. \tag{5.2.1}$$

Here, f_n are the **microscopic eigenvalues of the quantity.** Such mean values have a macroscopic meaning and are called **thermodynamic quantities.**

Naturally, deviations from the mean values of the quantities characterizing a body, i.e. **fluctuations**, always take place in the body. The **relative fluctuation** of an additive

quantity f is defined as

$$\frac{\sqrt{\langle (f-\langle f\rangle)^2\rangle}}{\langle f\rangle}. \tag{5.2.2}$$

A quantity is called **additive** if its value for the entire body is equal to the sum of its values for the subsystems constituting the body. For example, the energy of a body is equal to the sum of the energies of the subsystems to a very high degree of accuracy since the internal energy of the subsystem is much higher than the energy of interaction between the subsystems themselves. This follows from the fact that the relative number of particles on the surface of a subsystem (compared to the total number of particles in the subsystem) sharply decreases with increasing volume of the subsystem.

The radicand in (5.2.2) can be expanded as follows:

$$\langle (f - \langle f\rangle)^2\rangle = \langle [f^2 - 2f\langle f\rangle + (\langle f\rangle)^2]\rangle$$
$$= \langle f^2\rangle - 2\langle f\rangle\langle f\rangle + (\langle f\rangle)^2,$$

or

$$\langle (f - \langle f\rangle)^2\rangle = \langle f^2\rangle - (\langle f\rangle)^2.$$

Let us denote the deviation from the mean by $\Delta f = f - \langle f\rangle$. Obviously, $\langle \Delta f\rangle = 0$.

Since the number of subsystems in a body can be assumed to be proportional to the number N of particles constituting the body, it is obvious that $\langle f\rangle \propto N$ and $\langle (\Delta f)^2\rangle \propto N$ for the body as a whole. This leads to the following expression for the relative fluctuation of a macroscopic quantity characterizing a body:

$$\frac{\sqrt{\langle (\Delta f)^2\rangle}}{\langle f\rangle} \sim \frac{1}{\sqrt{N}}. \tag{5.2.3}$$

Thus, *the larger (in terms of the number of particles) and the more complex a body, the smaller the deviation of its thermodynamic quantities from the mean.* It was mentioned earlier that solids have a number density of the order of 10^{28} particles/m^3, and the fluctuations of the quantities from the mean in them are of the order of 10^{-14}.

5.3. Statistical Analysis of the Gas Model

5.3.1. Computer Experiments. A "theoretical experiment" was set up on a computer to study the interaction of several hundred spheres undergoing elastic collisions. The initial conditions (coordinates and velocities) and the equations of motion were specified for each sphere. The **equilibrium velocity distribution** in the system set in just after the first few individual collisions (see below). This equilibrium distribution was preserved and practically remained unchanged upon a change in the initial conditions for the spheres.

The result obtained with the help of the computer provides a visual and purely mechanical substantiation of the methods of statistical physics. Here we do not consider the motion of each individual particle in detail and completely neglect the initial conditions. We are interested only in the possible distributions of various groups of particles over the values of some quantity.

One of the most important results obtained on the computer is the law of distribution of "molecules" (spheres) in the bulk of a body. The experiment was carried out as follows. The initial positions of the "molecules" and their velocities were fed into the computer. The computer was asked to numerically solve the equations of motion of the "molecules" for the following instants of time and to show on an oscillograph their positions at these instants. A movie camera was used to take the motion picture of the oscillograph screen. The time interval was chosen in such a way that several "molecular collisions" occurred between two successive frames. A box divided arbitrarily into two equal left and right parts was chosen as the model of the body.

How are the "molecules" distributed between both halves of the box? In the case of a single molecule *1*, these distributions are shown in Fig. 5.1. Various distributions for two "molecules" *1* and *2* are shown in Fig. 5.2. These simple results can also be obtained without resorting to the "theoretical experiment" on a computer. Indeed, it is sufficient to count the number of different ways in which numbered spheres can be arranged in both halves of the box. Since for each possible location of one molecule there are two different

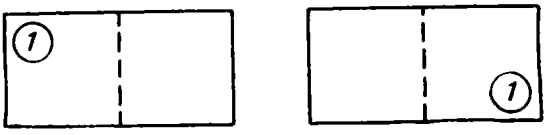


FIG. 5.1

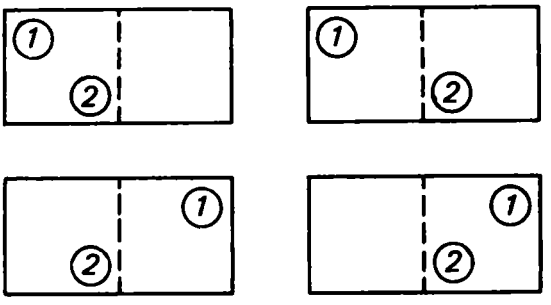


FIG. 5.2

ways in which the second molecule can be arranged, the total number of different distributions of two molecules is $2 \times 2 = 2^2 = 4$. In the general case of N molecules, the total number of possible different distributions in both halves of the box is 2^N.

5.3.2. Reversibility of Microscopic Processes in Time and Irreversibility of Macroscopic Processes. In the movie film of the oscillograph screen, each possible distribution of molecules has one frame corresponding to it. For N molecules, the number of frames is 2^N. Only one of these frames will show the distribution in which all N molecules are in the left half of the box. In other words, the probability of finding all N molecules in the left half of the box is given by

$$P_N = \frac{1}{2^N} .$$

Thus, for $N = 80$, the number of frames in the film is $2^{80} = 10^{24}$. For the film projected at the rate of 10^6 frames per second, it would take 10^{10} years to detect the frame in which all 80 molecules occupied the left half of the box.

This result can be interpreted as follows. Suppose that at the initial instant of time all the molecules are in the left

half of the box, separated from the right half by a shutter. Then the shutter is opened, and the movement and distribution of molecules is shot on a film. Obviously, the molecules will soon be evenly distributed between both halves of the box.

We project the film in the reverse direction (against time) so that the auditorium is not aware of this reversal. It can be well imagined that the spectators will burst into laughter at the end of the film. It will be clear to everybody that the operator has shown the film in the reverse order because the spectators will see an absolutely impossible situation in which the gas molecules move into the left half of the box on their own accord. If, however, the same procedure is adopted for a film involving two molecules only, nobody will be able to note whether the film is screened in the right direction or against time.

This means that *the processes occurring in macroscopic bodies are irreversible in time.* This is associated with the large number of particles involved in a process.

However, the temporal irreversibility of processes occurring in macroscopic bodies is of probabilistic nature. Small parts of a macroscopic system containing a few particles each may briefly experience nearly reversible processes or fluctuations (these were described quantitatively in Sec. 5.2). Although the fluctuation effects in macroscopic bodies are as a rule negligibly small, fluctuations do play a significant role in some phenomena (see, for example, Sec. 7.5.1).

5.4. Entropy

It was shown with the help of a classical molecular gas model that a macroscopic state can be reached in an isolated macroscopic system in a very short time. This state corresponds to the largest possible number of ways in which particles can be distributed over their microscopic states, and is called the **statistical equilibrium**. A body can remain in its equilibrium state for an indefinitely long time.

Going over to the more general "quantum" description, we introduce the concept of **entropy** S as the number $\Delta\Gamma$ of possible microscopic processes through which a given macro-

scopic state of a body can be realized. To be more precise, we can write

$$S = \ln \Delta\Gamma. \tag{5.4.1}$$

In particular, the lowest energy level of the body corresponds to $\Delta\Gamma = 1$ (see the last paragraph in Sec. 5.1), and hence to $S = 0$.

The probability w_n for different states of a body is proportional to $\Delta\Gamma$ which, in accordance with (5.4.1), is written as e^S. Consequently, we can write

$$w_n \propto e^S. \tag{5.4.2}$$

The most probable macroscopic state of a body (for a given energy) is the one corresponding to the largest number of microscopic distributions of particles. Hence, it is obvious that *entropy has its maximum value in the state of statistical equilibrium.*

If at a certain instant of time a closed system is not in a state of statistical equilibrium, processes of transition to more probable states take place in it. In other words, *the processes occurring in an isolated body are directed towards an increase in entropy.*

The **energy degradation law** can be formulated more precisely as follows: *entropy never decreases in a closed system*; *it either remains constant or increases.* In accordance with these two possibilities, all processes associated with a body are classified as **reversible** or **irreversible processes** respectively. Reversible processes can take place in both forward and backward directions since they do not involve a change in entropy. Irreversible processes cannot take place in the backward direction since this involves a decrease in entropy.

5.5. Temperature

From the definition of entropy, it can be concluded that entropy S is an increasing function of the internal energy E. This is illustrated in Fig. 5.3. The function $S(E)$ is represented in Fig. 5.4.

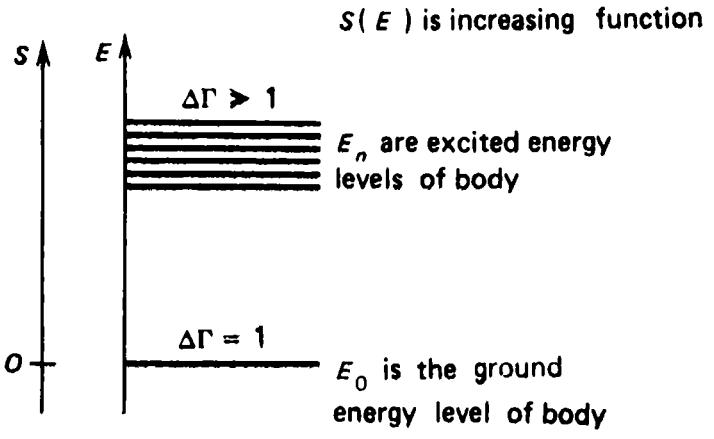


FIG. 5.3

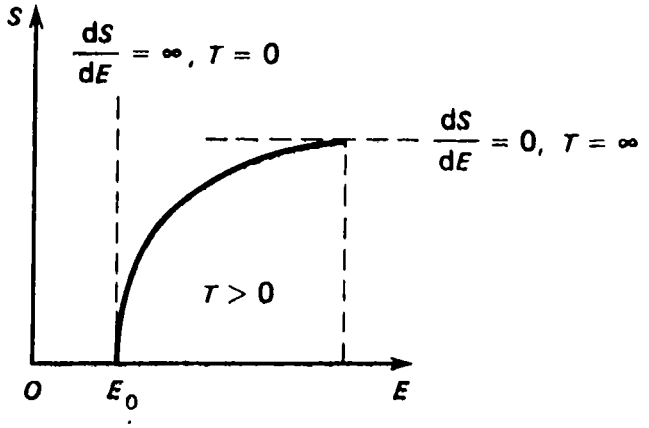


FIG. 5.4

The (thermodynamic) **temperature** T is defined as

$$\frac{dS}{dE} = \frac{1}{T}. \tag{5.5.1}$$

In this connection, it is essential to recall the mathematical addendum to Sec. 4.5 (on the concept of derivative and its geometrical interpretation).

Let us consider the basic properties of temperature. For this, we turn to Fig. 5.4 which shows that at $T = 0$, a body can exist only in the ground (lowest) quantum state. The temperature T may vary in the interval $0 \leqslant T < \infty$ and is always positive.

We shall show that *for two bodies in the state of statistical equilibrium, their temperatures* T_1 *and* T_2 *are identical.*

Suppose that both bodies form a single closed system with a constant energy E. In the state of statistical equilibrium, the entropy S of the entire system must have its maximum possible value (for the given energy E). Energy and entropy are additive quantities, i.e. $E = E_1 + E_2$ and $S = S_1(E_1) + S_2(E_2)$, where E_1, E_2 and S_1, S_2 denote the energy and entropy of each body respectively. Since $E_2 = E - E_1$, S is a function of just one variable E_1. The condition for maximum value of the function $S(E_1)$ is written in the form (see condition (4.5.6))

$$\frac{dS}{dE_1} = \frac{dS_1}{dE_1} + \frac{dS_2}{dE_2}\frac{dE_2}{dE_1} = 0.$$

We have applied the chain rule to the second term on the right-hand side. The derivative $dE_2/dE_1 = -1$ since E is constant. As a result, we get

$$\frac{dS}{dE_1} = \frac{dS_1}{dE_1} - \frac{dS_2}{dE_2} = 0,$$

i.e.

$$\frac{dS_1}{dE_1} = \frac{dS_2}{dE_2}.$$

This equality can be generalized to the case of an arbitrary number of bodies in statistical equilibrium. In other words (see Eq. (5.5.1)), *the temperature of a system in equilibrium is the same for all its parts.* Hence statistical equilibrium is also called **thermal equilibrium.**

Suppose that a closed system of two bodies with temperatures T_1 and T_2 is not in thermal equilibrium. In this case, the temperatures of these bodies are different, say, $T_2 > T_1$. The bodies undergo a transition to a common equilibrium state with a gradual equalization of temperature throughout the system. The entropy $S = S_1 + S_2$ of the system increases in this case. In mathematical terms, it means that the time derivative of entropy is a positive quantity, i.e.

$$\frac{dS}{dt} = \frac{dS_1}{dt} + \frac{dS_2}{dt} = \frac{dS_1}{dE_1}\frac{dE_1}{dt} + \frac{dS_2}{dE_2}\frac{dE_2}{dt} > 0.$$

Since the total energy $E = E_1 + E_2$ is conserved, we get

$$\frac{dE}{dt} = \frac{dE_1}{dt} + \frac{dE_2}{dt} = 0. \qquad (*)$$

Consequently,

$$\frac{dS}{dt} = \left(\frac{dS_1}{dE_1} - \frac{dS_2}{dE_2}\right)\frac{dE_1}{dt} = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\frac{dE_1}{dt} > 0.$$

However, in accordance with the stipulation made above, $T_2 > T_1$. Hence we obtain

$$\frac{dE_1}{dt} > 0.$$

From (*), we find that the derivative

$$\frac{dE_2}{dt} < 0.$$

This means that the energy of the first body increases, while that of the second body decreases. Thus, this property of temperature can briefly be described as follows: *an energy flow is created between bodies with different temperatures, the energy flowing from the body with a higher temperature to the body with a lower temperature.*

It should be observed that since entropy is a dimensionless quantity, temperature has the dimensions of energy in accordance with Eq. (5.5.1). In order to use Kelvin's temperature scale, we must employ the conversion factor

$$k = 1.38 \times 10^{-23} \text{ J/K},$$

called the **Boltzmann constant**. The transition from the temperature in the energy units (T_E) to the temperature in kelvins (T_K) is made with the help of the rule

$$T_E \to kT_K.$$

In future, we shall omit the subscripts in the notation for temperature and explain the meaning of a particular notation when necessary.

Although temperature, like entropy, is of statistical nature, let us try to find its "microscopic" meaning. We present (5.5.1) in terms of finite differences:

$$\frac{1}{T} = \frac{\Delta S}{\Delta E} = \frac{S - S_0}{E - E_0}, \qquad (5.5.2)$$

where E_0 is the energy of the ground quantum state of a body, and $S_0 = 0$ is the entropy of the body in the ground state. In view of the additive nature of entropy, $S \propto N$,

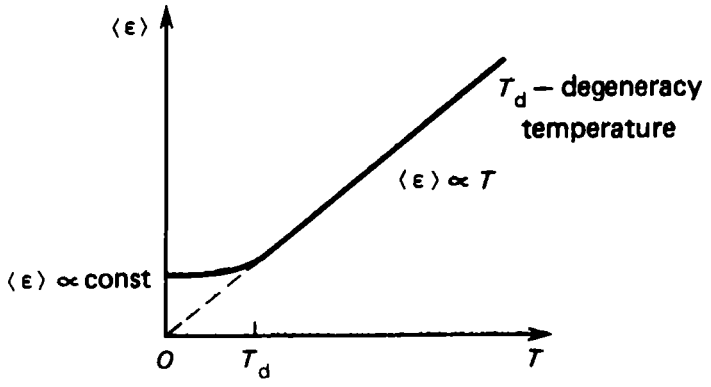


FIG. 5.5

where N is the total number of particles in the body. Consequently, we can write (5.5.2) in the form

$$\frac{1}{T} = \frac{S}{E-E_0} \propto \frac{N}{E-E_0} ,$$

or

$$T \propto \frac{E-E_0}{N} = \langle \varepsilon \rangle . \tag{5.5.3}$$

In the last equation, $\langle \varepsilon \rangle$ is the mean energy of an individual particle in the body. For $E \gg E_0$ (classical region of the energy states of the body), $\langle \varepsilon \rangle$ is proportional to T. However, for $E = E_n \sim E_0$ (quantum region), $\langle \varepsilon \rangle$ assumes a constant value which is independent of the temperature.

The results of the above analysis are represented in Fig. 5.5.

It was mentioned above that the temperature of a macroscopic body as a whole can only be positive. This can also be verified as follows. In view of the additive nature of entropy, we can write

$$S_i \left(E_i - \frac{\mathbf{p}_i^2}{2M_i} \right) , \tag{5.5.4}$$

where the subscript i indicates the parameters describing small but macroscopic regions of the body (also called its subsystems), E_i, $\mathbf{p}_i$, and M_i describe the total energy, momentum, and mass of individual subsystems respectively. The argument of the entropy S_i of the subsystems describes their internal energy. If a body could have a temperature $T < 0$, then in accordance with (5.5.1), its entropy would

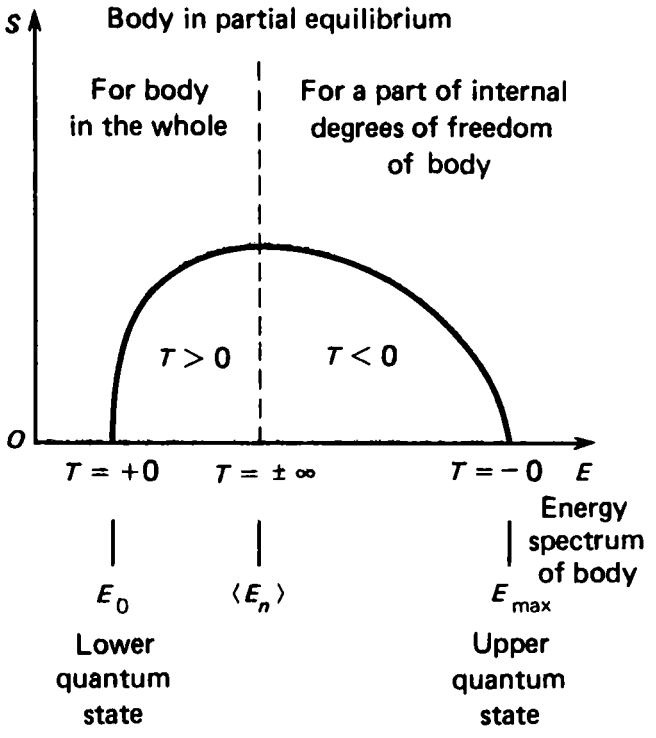


FIG. 5.6

increase with decreasing internal energy. However, a decrease in the argument of the entropy is possible only through an increase in the kinetic energy of the subsystems. This makes the existence of stable macroscopic bodies impossible.

If a body is not completely in equilibrium, it may so happen that a part of the internal degrees of freedom of the body (not connected with its translational motion) corresponds to negative temperatures. Of course, the body as a whole has a positive temperature. This situation is represented in general form in Fig. 5.6. Some particular cases will be considered below.

5.6. Equilibrium Distribution of Particles in a Body

Let us turn to the basic problem of statistical physics formulated in Sec. 5.1, i.e. that of determining the law of distribution of probabilities w_n of equilibrium macroscopic systems existing in different quantum states with energies E_n.

Let us consider a small part of a certain body (its subsystem). Suppose that this part occupies a quantum state with energy E_0. The energies of the neighbouring quantum states of the subsystem can be represented by E_n. The entropy S is a function of $(E_0 - E_n)$. Since the energy levels of a macroscopic body are separated by quite small distances, we can write

$$S(E_0 - E_n) \approx S(E_0) - E_n \frac{dS}{dE}. \tag{5.6.1}$$

The gap between energy levels can be estimated by using the quantum uncertainty principle. Denoting by $D(E)$ the mean separation between energy levels, we obtain

$$D(E) \sim \frac{(\Delta p)^2}{m} \sim \frac{\hbar^2}{mL^2}. \tag{5.6.2}$$

In this relation, we have considered a particle of mass m in a "box" of linear dimensions L. For example, in the case of an electron in a "potential well" with $L \simeq 10^{-2}$ m $=$ 1 cm, we find that $D(E) \sim 10^{-34}$ J. For the sake of comparison, it may be recalled that the energy levels of an atom for which $L \simeq 10^{-10}$ m are separated by about 10^{-18} J.

It should be borne in mind that *macroscopic bodies cannot exist in a strictly stationary quantum state with a definite energy value* $E_n \equiv E$. This is so because howsoever small the external influence on the body, it is nevertheless much larger than the mean separation $D(E)$ between the quantum levels in the energy spectrum of the body. In other words, the energy of the body is always "spread" by an amount ΔE about the value E. This energy gap ΔE embraces a large number of energy levels which are almost continuous in the spectrum.

Let us substitute expression (5.6.1) for S into (5.4.2). Taking into account definition (5.5.1), we obtain

$$w_n \propto e^{-E_n/(kT)}. \tag{5.6.3}$$

Here, we must take into consideration the fact that the sum of the probabilities of all possible states of the system must be unity (after all, the system *must* be in *some* state). In-

troducing the normalization constant A into (5.6.3), we finally get

$$w_n = A\,e^{-E_n/(kT)}. \tag{5.6.4}$$

This relation is called **Gibbs distribution.** It is applicable not only to subsystems in statistical equilibrium but also to the equilibrium bodies themselves since a body can be treated as a part of a larger system, viz. the ambient.

Let us apply the Gibbs distribution to the "population" of particles on the quantum energy levels of a body. The body is chosen in the form of a "gas" containing N particles (molecules). The interaction between the particles will be neglected. We introduce the mean occupation numbers n_i of different molecular quantum levels of energy E_i. This gives

$$N = n_0 + n_1 + n_2 + \ldots + n_m$$

and

$$E = n_0E_0 + n_1E_1 + n_2E_2 + \ldots + n_mE_m,$$

where E is the total energy of the body which is conserved.

The probability of finding a certain number of particles on the ith quantum level is given by

$$\frac{n_i}{N} \propto e^{-E_i/(kT)}.$$

The ratio of the mean occupation numbers for different levels is written as

$$\frac{n_1}{n_0} = e^{-(E_1-E_0)/(kT)}. \tag{5.6.5}$$

By way of an example, we have considered in this equation the ratio of the population of the first excited state to that of the ground state. It can be shown that if the sequence of energy values E_i of the levels forms an arithmetic progression, the corresponding occupation numbers n_i form a geometric progression. The general form of this dependence is called **Boltzmann distribution** and is shown graphically in Fig. 5.7.

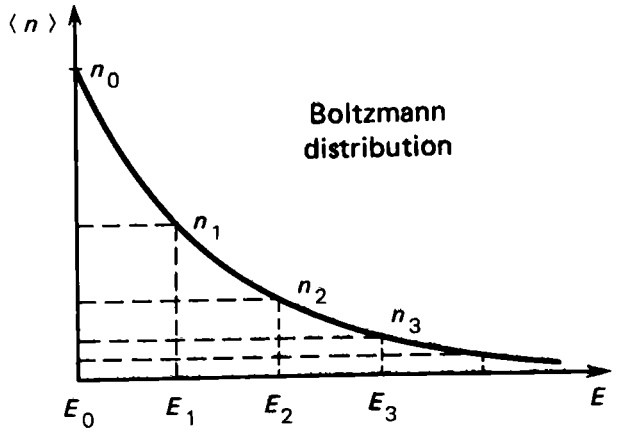


FIG. 5.7

It is interesting to analyze the temperature dependence of the population of levels in a system. For this purpose, we take the logarithms of both sides in (5.6.5). This gives

$$T = \frac{E_1 - E_0}{k \ln (n_0/n_1)} . \tag{5.6.6}$$

This expression for the temperature could also be obtained directly from Eq. (5.5.1) by taking into account Eq. (5.4.1). Let us analyze formula (5.6.6). Since $E_1 - E_0 > 0$, we obtain

(i) $T = 0$ for $n_0 \neq 0$ and $n_1 = 0$;
(ii) $T > 0$ for $n_0 > n_1 \neq 0$;
(iii) $T = +\infty$ for $n_1 = n_0$.

If, however, the population of levels in a system (to be more precise, in a subsystem of a certain larger system) is such that $n_1 > n_0$, it means that $T < 0$. Such a situation corresponds to a negative temperature. In particular, if $n_1 \neq 0$ and $n_0 = 0$, we obtain $T = -0$ (Fig. 5.8).

It can be seen from Fig. 5.8 that negative temperatures correspond to levels higher than $T = +\infty$. This is in conformity with the fact that energy is spontaneously transferred from parts of the system with $T < 0$ to parts with $T > 0$. Let us compare this figure with Fig. 5.6, where we have considered by way of an example the possible behaviour of the subsystem of atomic magnetic moments for a crystal lattice under the assumption that they interact weakly with the lattice.

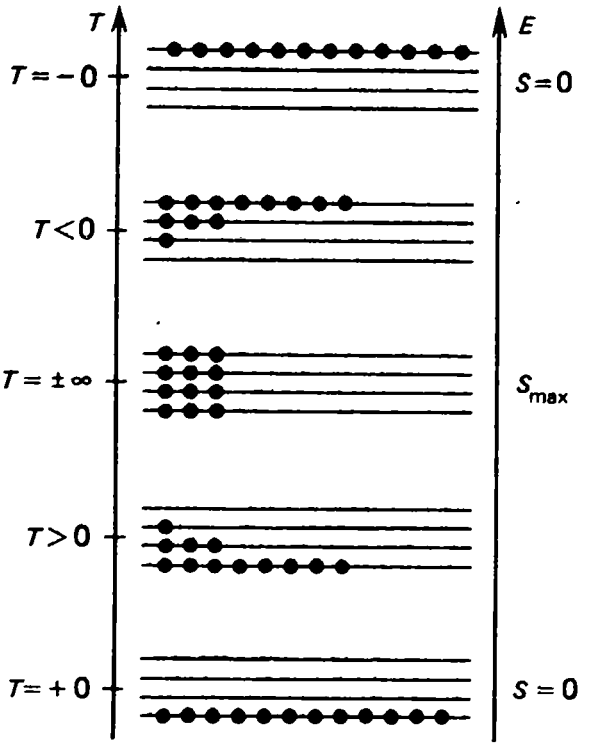


FIG. 5.8

The subsystem may not be in complete statistical equilibrium with the crystal (e.g. when it is excited by an external magnetic field). The aggregate of magnetic moments has finite energies (of the order of the interaction energy of magnetic moments) and a finite number of levels. The magnetic moments themselves are fixed at the lattice sites and can only be oriented in a magnetic field by changing their energy in the process.

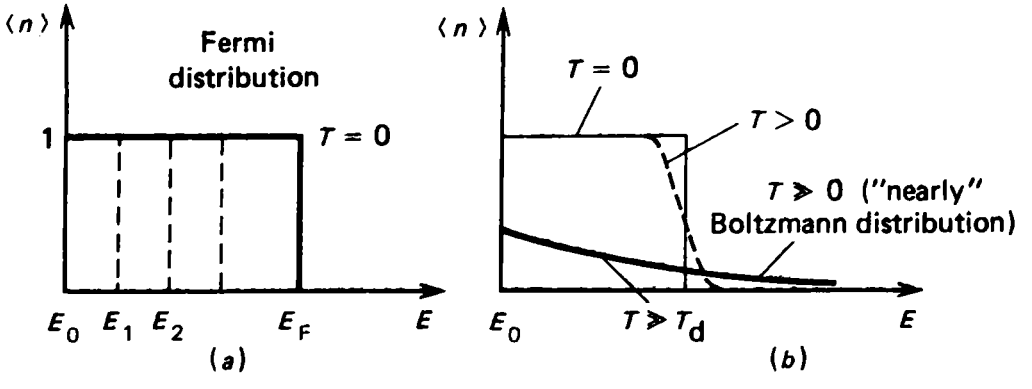


FIG. 5.9

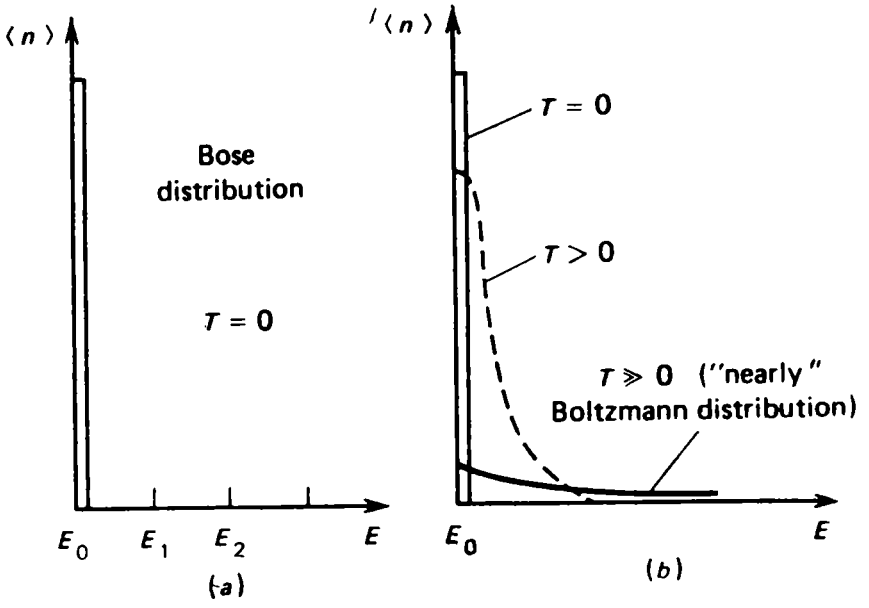


FIG. 5.10

It is well known (see Sec. 4.6.2) that a system of identical particles obeys special rules for populating quantum states by these particles. If the particles have half-integral spins, each quantum state of the system can accommodate not more than one particle. The resulting statistics of the particles is called **Fermi distribution** (Fig. 5.9*a*). The nature of this distribution for different temperatures is shown in Fig. 5.9*b*. If, however, the particles have integral spins, there are no restrictions on the occupation numbers of quantum states (Fig. 5.10*a*). The evolution of this distribution **(Bose distribution)** for different temperatures is shown in Fig. 5.10*b* (see Sec. 5.8.2 on details of quantum statistics).

Later, we shall study the applications of Fermi and Bose distributions by considering specific examples. The existence of these statistics reveals quite unusual properties of matter.

5.7. Thermodynamic Relations

Knowing the law of distribution (5.6.4), we can use formula (5.2.1) to determine the values of thermodynamic quantities (see Sec. 5.2). However, thermodynamic quantities may be connected through relations that are applicable to all bodies irrespective of their atomic structure.

The general method used for determining these relations is called **thermodynamics**. Thermodynamics may be constructed axiomatically without any reference to statistics. The basic concepts of thermodynamics are covered by three laws of thermodynamics, viz. (i) the energy conservation law; (ii) the energy degradation law (the very concept of entropy is introduced phenomenologically in thermodynamics); (iii) the vanishing of entropy at absolute zero.

These laws of thermodynamics do not present anything new. We have discussed these statements statistically while studying the behaviour of bodies as aggregates of particles. Although physical statistics alone can lead to concepts of thermodynamics, only the latter has a right to independent existence since it is based on general macroscopic relations disregarding the microscopic structure of bodies.

If we are not quite familiar with the atomic structure of a macroscopic object in which complex processes take place, thermodynamic relations seem to be the only course (besides experiment) for studying the thermal equilibrium properties of such objects.

Let us derive these thermodynamic relations. The change dE in the internal energy of a body must be the sum of the "thermal" term equal to $T\,dS$ (in accordance with Eq. (5.5.1)) and the "mechanical" term equal to $-p\,dV$. The expression for the latter term is obtained as a result of comparison with mechanics:

$$\left(F = -\frac{dU}{dr}\right) \rightarrow \left(p = -\frac{dE}{dV}\right),$$

where F is the force, p the pressure, and U the potential energy of interaction between particles, which is a function of the distance r between them. Consequently, the total variation of the internal energy of a body is defined as

$$dE = T\,dS - p\,dV. \tag{5.7.1}$$

This equation contains two completely different quantities, one of which, viz.

$$dQ = T\,dS \tag{5.7.2}$$

is called the **amount of heat**, and the other, viz.

$$dA = -p\,dV \tag{5.7.3}$$

is called **work.** Equation (5.7.1) can be written alternatively in the form

$$dE = dQ + dA. \tag{5.7.4}$$

In addition to Eqs. (5.7.1) and (5.7.4) which express the energy conservation law, there also exists an important relation which arises from the concept of **maximum work.** Suppose that a nonequilibrium system composed of several bodies is thermally insulated, i.e. obeys the relation $dQ = 0$. As the system attains thermal equilibrium, it does work over external objects. This equilibrium may be attained in different ways, and the final energy and entropy of the system will be different in these cases. The total work done by the nonequilibrium system (its total volume remaining unchanged) depends on the way in which the thermal equilibrium is attained. Let us find the mode of equilibrium attainment for which the maximum amount of work is done.

Let E_{in} be the initial energy of a nonequilibrium system, and $E(S)$ the energy as a function of entropy in the final equilibrium state. Since $dQ = 0$, we can write the following expression for work in accordance with (5.7.4):

$$|A| = E_{in} - E(S).$$

By assumption, the work done by the system is negative. The derivative of $|A|$ with respect to the entropy of the final state is written as

$$\frac{d|A|}{dS} = -\frac{dE}{dS} = -T.$$

This derivative is negative, i.e. $|A|$ decreases with increasing S. If the entropy were to decrease, the work would increase. Hence the maximum possible work is done by a nonequilibrium system when its entropy does not change, i.e. when the process of equilibrium attainment is reversible (see the end of Sec. 5.4).

Let us determine the quantity $|A|_{max}$ during the exchange of energy between two bodies *1* and *2* with temperatures T_1 and T_2. We assume that $T_2 > T_1$, and hence call bodies *2* and *1* a **heater** and a **cooler** respectively. We introduce an intermediate **working substance,** otherwise two bodies with different temperatures will exchange energy irreversibly upon coming in contact and will not do any

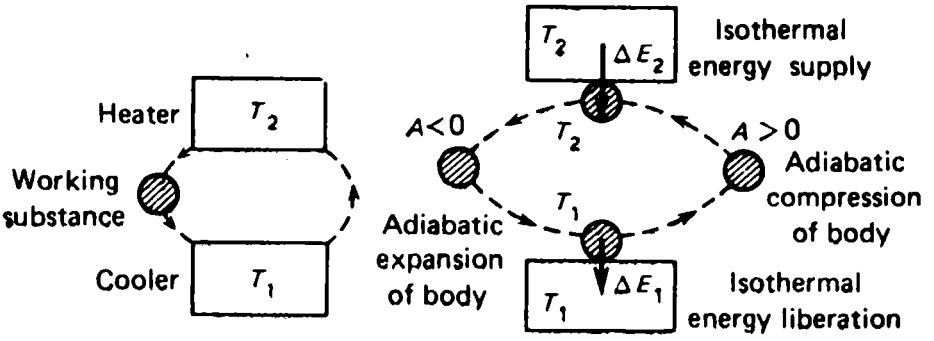


FIG. 5.11

work (see Sec. 5.5). Suppose that the working substance undergoes a reversible cyclic process (Fig. 5.11). Since the working substance returns to its initial position at the end of a cyclic process, we shall not consider it during the quantitative analysis for determining the amount of work. We assume that

$$-\Delta E_2 = -T_2 \Delta S_2$$

is the energy given away by the hot body, and

$$\Delta E_1 = T_1 \Delta S_1$$

is the energy received by the cold body. Since the process is reversible and the entropy of the system is constant, we get

$$\Delta S_1 + \Delta S_2 = 0,$$

or

$$\Delta S_1 = -\Delta S_2.$$

The total work done is equal to the decrease in the total energy of both bodies and is defined as

$$|\Delta A|_{\max} = -\Delta E_1 - \Delta E_2 = -T_1 \Delta S_1 - T_2 \Delta S_2$$
$$= -(T_2 - T_1) \Delta S_2,$$

or

$$|\Delta A|_{\max} = \frac{T_2 - T_1}{T_2} |\Delta E_2|. \qquad (5.7.5)$$

We introduce the **efficiency** η defined as the ratio of the work $|\Delta A|$ done to the energy expenditure $|\Delta E_2|$. This gives

$$\eta_{max} = \frac{T_2 - T_1}{T_2} . \qquad (5.7.6)$$

Formulas (5.7.5) and (5.7.6) essentially reflect the statement that *it is impossible to do work on account of the energy of bodies in thermal equilibrium* (i.e. when $T_2 = T_1$). These formulas are important not only from a theoretical point of view but also because of their practical applications.

5.8. Ideal Gas

5.8.1. Matter and Its States. Let us consider once again the macroscopic bodies as aggregates of particles. Large aggregates of particles are also called **matter**. Matter may be in **solid, liquid** or **gaseous state.**

Matter in different states has different microstructure and thermal motion. In gases, particles perform a random translational motion. In crystalline solids, there is no translational motion of the particles which perform a vibrational motion about certain fixed equilibrium positions. In liquids, the motion of particles is a combination of translational and vibrational motions.

5.8.2. Classical and Quantum Ideal Gases. Matter in the gaseous state is an aggregate of many weakly interacting particles. Since the kinetic energy of gas particles is much higher than their interaction energy, the latter is often neglected during an analysis of the properties of gases. Such a model is called an **ideal gas.**

Examples of the systems treated as ideal gases are (i) atomic and molecular gases at normal temperatures and pressures; (ii) dense gases at high temperatures; (iii) a "gas" of free electrons in metals; (iv) the "photon gas" of free radiation in a cavity.

As a many-particle system, a gas obeys classical laws for such values of density and temperature for which the quantum-mechanical probability wavelength λ corresponding to the thermal motion of particles is small in comparison with the mean separation l between the particles, $\lambda \ll l$ (see

Sec. 4.7.1 and Fig. 4.10). If the value of λ is comparable with l ($\lambda \sim l$), the gas obeys quantum laws.

For $\lambda \ll l$, only a fraction of the degrees of freedom of molecules or atoms is quasi-classical in nature. The translational motion of molecules characterized by their momentum is quasi-classical (the equilibrium momentum distribution of molecules is represented by Maxwell distribution, see below). However, the internal state of molecules is naturally considered quantum mechanically (the equilibrium distribution of molecules over quantum states is described by Boltzmann distribution, see Fig. 5.7).

When $\lambda \sim l$, the equilibrium distribution of particles over different quantum states (for a system of identical particles) is represented by Fermi or Bose distribution (see Figs. 5.9 and 5.10), depending on the spin of the particles. For $\lambda \ll l$, both Fermi and Bose distributions are transformed into Boltzmann distribution.

It should be observed that a more general criterion for determining whether a gas is a classical or a quantum system is the ratio of the number of occupied quantum states to the total number of quantum states. For a classical gas, this ratio is much smaller than unity, while for a quantum gas, it is comparable with unity.

Atomic and molecular gases are classical gases. Gases of light elementary particles with a relatively high density can be treated as quantum ideal gases.

Let us consider the energy spectra of quantum ideal gases. From a physical point of view, the spectra of systems of identical particles are the most interesting. It was mentioned in Sec. 4.6.2 that certain restrictions are imposed on the possible states of such systems, and only two types of systems, viz. Bose- and Fermi systems, can exist.

The difference in the properties of ideal Bose- and Fermi gases can be judged from an analysis of their energy spectra at absolute zero ($T = 0$ K). Since for a Bose gas at $T = 0$ K the energy spectrum of the system is essentially represented by a single ground level with zero energy, excited states practically do not appear. All particles of the system are located on this lowest ground level since an arbitrary number of particles can exist simultaneously in the same quantum state for a Bose system (see Fig. 5.10).

In an ideal Fermi gas at $T = 0$ K, the energy spectrum

is a quasi-continuous band of levels starting from the zero level and terminating at a certain limiting level E_F whose energy depends on the number of particles in the system. Such a nature of the spectrum is due to the fact that a quantum state of a Fermi system cannot contain simultaneously more than one particle (see Fig. 5.9).

An analysis of the spectra of quantum ideal gases leads to the conclusion that the total energy of a Fermi gas at $T = 0$ K is nonzero, while for a Bose gas, the total energy is zero.

5.8.3. Equation of State for an Ideal Gas. The basic quantity characterizing a macroscopic system is its energy E. Since energy is an additive quantity, E is a function of the volume V. It should be observed that the pressure $p \approx E/V$. For an ideal gas of N atoms in a random translational motion, we can write the following expressions for the energy and pressure:

$$E = \sum_{i=1}^{N} \frac{m}{2} \langle v_i^2 \rangle = N \langle \varepsilon \rangle \tag{5.8.1}$$

and

$$p = \frac{N}{V} \langle \varepsilon \rangle = \frac{NkT}{V}. \tag{5.8.2}$$

While deriving Eq. (5.8.2), we have used expression (5.5.3). Equation (5.8.2) written in the form

$$pV = NkT \tag{5.8.3}$$

is called the **equation of state for a classical ideal gas.** This relation connects the pressure, volume, and temperature of a gas. Since Eq. (5.8.3) does not contain any quantities characterizing the nature of a specific gas, it can be applied to any classical ideal gas. This is in accord with the fact that while deriving Eq. (5.8.3), we neglected the interaction of atoms, thus depriving a gas of its "individuality".

Two different ideal gases having the same values of p, V, and T contain equal numbers of atoms. For example, one cubic metre of a gas under normal conditions (viz. $T = 273$ K and $p = 1$ atm $= 1.01 \times 10^5$ Pa) contains

$$\frac{N}{V} = \frac{p}{kT}$$

atoms. Substituting the numerical values, we obtain

$$\frac{N}{V} = \frac{1.01 \times 10^5}{1.38 \times 10^{-23} \times 273} \text{ atoms/m}^3 = 2.69 \times 10^{25} \text{ atoms/m}^3.$$

The number of atoms in one mole of a gas, i.e. $V = 22.4 \times 10^{-3}$ m^3, is

$$N_A = 6.02 \times 10^{23} \text{ mol}^{-1}.$$

This number is called **Avogadro's number.**

Let us write the equation of state (5.8.3) for 1 mole:

$$pV = RT, \tag{5.8.4}$$

where $R = N_A k$ is called the **molar gas constant:**

$$R = 8.3 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}.$$

We shall now derive the classical **Maxwell distribution** mentioned in the previous section. We consider the equilibrium velocity distribution of gas atoms in a nearly free translational motion. For this purpose, it is sufficient to replace E_n in the Gibbs distribution (5.6.3) by $mv^2/2$ since the translational motion of an atom can be described quite accurately by the laws of motion of classical mechanics (see Fig. 4.10). Consequently,

$$w(v) \propto e^{-mv^2/(2kT)}. \tag{5.8.5}$$

In order to obtain the distribution of the mean numbers $\langle n(v) \rangle$ of particles over the magnitudes of their velocities v (having a spread dv), it is sufficient to multiply (5.8.5) by the volume of a spherical layer $4\pi v^2\, dv$ (in the "velocity space"). This gives

$$\langle n(v) \rangle \equiv \frac{dN}{dv} = \text{const}\, v^2 e^{-mv^2/(2kT)}. \tag{5.8.6}$$

This function is graphically represented in Fig. 5.12. The peak of the curve lies at $v = \sqrt{2kT/m}$. This velocity is slightly lower than the root-mean-square velocity $\sqrt{\langle v^2 \rangle}$, which can be calculated from the Maxwell distribution (5.8.5) and is equal to $\sqrt{3kT/m}$. Hence we obtain the following expression for the mean kinetic energy of an atom:

$$\frac{m}{2} \langle v^2 \rangle = \frac{3}{2} kT. \tag{5.8.7}$$

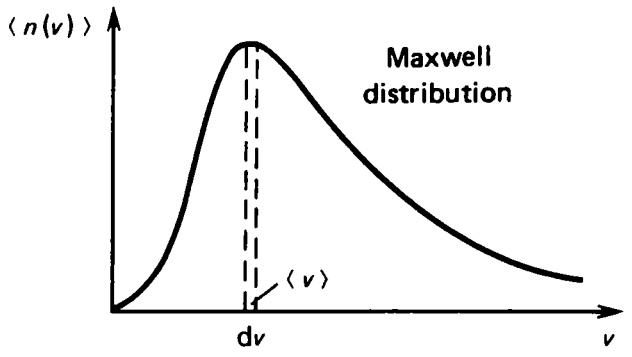


FIG. 5.12

A comparison of this result with Fig. 5.5 shows that the range of applicability of (5.8.7) is limited from below by the **degeneracy temperatures** T_d. At these temperatures, quantum effects appear in an ideal gas, and the equation of state (5.8.2) for a gas must be refined.

Let us derive an expression for the degeneracy temperature T_d. It was mentioned above that the quantum effects become significant in a system for

$$\lambda \sim l, \tag{5.8.8}$$

where λ is the quantum-mechanical wavelength of a particle, and l is the mean separation between particles. Taking Eq. (4.3.1) into consideration, we can write Eq. (5.8.8) in the form

$$\frac{\hbar}{p} \sim \left(\frac{N}{V}\right)^{-1/3}. \tag{5.8.9}$$

Here, p is the particle momentum, and (N/V) is the particle density. Assuming, in accordance with (5.8.7), that the mean kinetic energy of a particle in the system is given by

$$\frac{p^2}{m} \sim kT,$$

we obtain from (5.8.9)

$$T_d \sim \frac{\hbar^2}{mk}\left(\frac{N}{V}\right)^{2/3}. \tag{5.8.10}$$

Thus, the degeneracy temperature is determined by the mass and density of particles. A substitution of numerical values for all atomic (and molecular) gases shows that quantum effects are manifested only for densities at which gases are

actually in the condensed state (liquids or crystals). The situation is quite different for gases of light elementary particles (see below).

5.8.4. Heat Capacity of an Ideal Gas. The **heat capacity** C is defined as

$$C = \frac{dQ}{dT}. \tag{5.8.11}$$

This quantity is characterized by the amount of heat which must be transferred to a body to raise its temperature by one kelvin. Obviously, such a definition of heat capacity is ambiguous. Indeed, the absorbed energy is used not only for increasing the temperature of the body but also, say, for its expansion. In other words, while defining heat capacity, we must specify the conditions under which the body is heated. In this connection, two types of heat capacities are considered, viz. C_V, the heat capacity at constant volume V of the body, and C_p, the heat capacity at constant external pressure p.

For a body of constant volume, $dV = 0$, and hence, according to Eqs. (5.7.1) and (5.7.2), $dQ = dE$. In other words, the entire heat received by the body is spent on changing its internal energy. Therefore, we can write

$$C_V = \left(\frac{dE}{dT}\right)_V. \tag{5.8.12}$$

The subscript V on the derivative means that the differentiation is carried out at constant volume V (after all, the internal energy E is a function of a large number of variables including V and T).

When a body is heated under the condition that the external pressure p remains constant, which is the most natural case in actual practice, we obtain

$$dQ = dE + p\,dV. \tag{5.8.13}$$

In this case, the absorbed heat is spent not only for increasing the internal energy of the body but also for the work of its expansion. Equation (5.8.13) can also be written in the form

$$dQ = d(E + pV) \equiv dH. \tag{5.8.14}$$

Consequently, the heat capacity is given by

$$C_p = \left(\frac{dH}{dT}\right)_p, \tag{5.8.15}$$

where H is called **enthalpy (thermal function).**

Obviously, the heat capacity C_p is always larger than C_V:

$$C_p > C_V. \tag{5.8.16}$$

Before considering the heat capacity of ideal gases, we observe that since the interaction between particles in ideal gases is weak and therefore neglected in computations, the internal energy E for a given mass of the gas is a function of only T, and not of V and p.

The enthalpy of one mole of a gas can be written in the form

$$H \equiv E + pV = E + RT, \tag{5.8.17}$$

where we have used the equation of state (5.8.4) for an ideal gas. Differentiating (5.8.17) with respect to temperature and taking into account definitions (5.8.12) and (5.8.15) of heat capacity, we obtain

$$C_p = C_V + R. \tag{5.8.18}$$

This is the simple relation connecting the two heat capacities of an ideal gas.

Let us determine the heat capacity of a *monatomic gas* (inert gases are an example of an atomic gas). Since the internal energy of such gases is described by Eq. (5.8.1), taking (5.8.7) into consideration, we can write

$$E = \frac{3}{2} kTN_A = \frac{3}{2} RT. \tag{5.8.19}$$

According to Eqs. (5.8.12) and (5.8.18), the heat capacities at constant volume and constant pressure are given by the expressions

$$C_V = \frac{3}{2} R, \quad C_p = \frac{5}{2} R. \tag{5.8.20}$$

Let us calculate the heat capacity of a *diatomic gas* (i.e. the gas whose molecules consist of two atoms). For this purpose, we must turn to the structure of these molecules. It is well known from Sec. 4.10 that the energy spectrum of a di-

atomic molecule consists of several branches, viz. E_{elec}, E_{vib}, and E_{rot}. Only the excited states of a spectral branch can participate in thermal motion. Hence the thermal dynamics of a system at temperature T mainly "includes" the branches of quantum states for which $T \gtrsim \Delta E_i$, ΔE_i being the separation between the corresponding energy levels. The order of magnitude (in eV) of the energy gap in various branches of the energy spectrum of a molecule was described in Sec. 4.10. In order to use these data, we observe that

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ J} = \frac{1.6 \times 10^{-19}}{1.38 \times 10^{-23}} \text{ K} = 11\,600 \text{ K}.$$

In this case, $\Delta E_{elec} \sim kT$ corresponds to a temperature of several tens of thousand kelvins, ΔE_{vib}, to a temperature of over one thousand kelvins, and ΔE_{rot}, to a temperature of about 10 K. The energy of translational motion of a molecule as a whole changes continuously. Thus, for example, if we consider a diatomic gas (hydrogen, nitrogen, oxygen) at room temperature, only the translational and rotational "degrees of freedom" of molecules will actually participate in thermal motion.

An isolated atom has only three degrees of freedom related with the translational motion of the atom as a whole. A molecule consisting of two atoms has three degrees of freedom for its translational motion (Fig. 5.13*a*) and two more degrees of freedom for rotational motion (the vibrational motion of atoms in the molecule is not important at this stage) (Fig. 5.13*b*). There is no sense in considering the rotation of molecule about the X-axis.

If the temperature of the system is $T \gg \Delta E_n$, where ΔE_n are the characteristic energy gaps in its energy spectrum, then (i) the properties of the system are close to classical; (ii) in view of the disordered nature of thermal motion, none of the types of motion of a complex particle (say, a molecule) predominates over the other types; (iii) in statistical equilibrium, each translational or rotational degree of freedom of a complex particle has the same mean kinetic energy $kT/2$ corresponding to it, while each vibrational degree of freedom has an energy kT corresponding to it.

Without taking into account the vibrational degrees of freedom, we obtain the following expressions for the heat

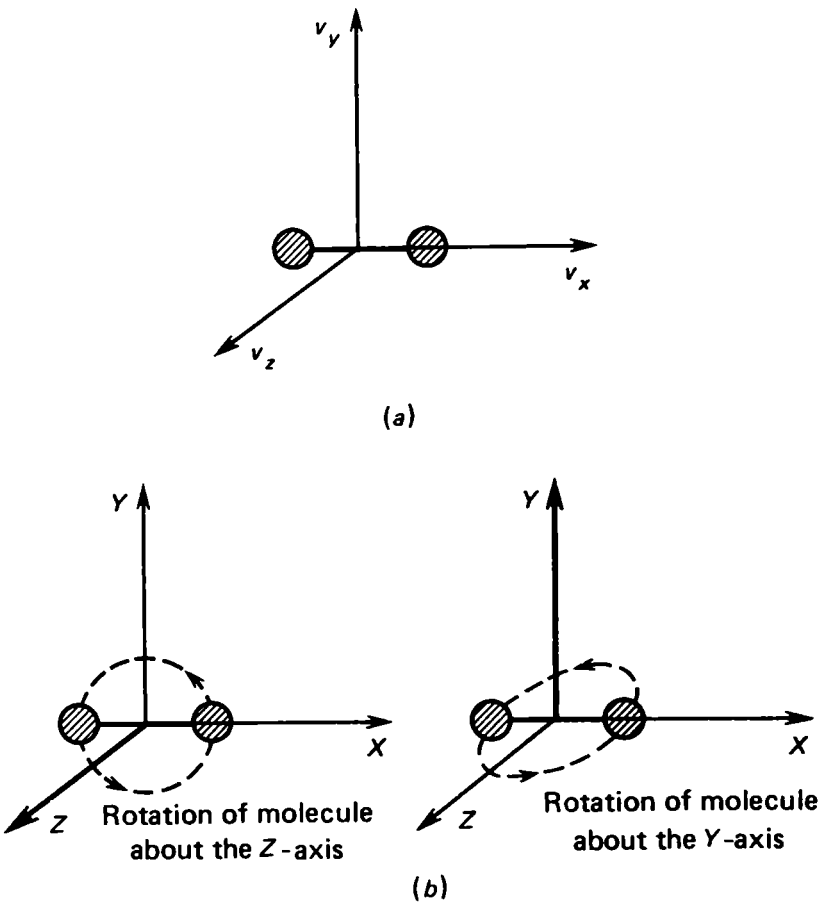


FIG. 5.13

capacities of a diatomic gas (at room temperature):

$$C_V = \frac{5}{2} R, \quad C_p = \frac{7}{2} R. \tag{5.8.21}$$

As the temperature increases ($T \sim 10^3$ K), the vibrational degrees of freedom will appear in the thermal motion, and the heat capacity of the gas will accordingly change.

5.8.5. Reversible Thermal Processes. If a process occurs in such a way that the system is in a state of thermal equilibrium at each instant of time and its entropy does not change during the process, the process is called reversible (see the end of Sec. 5.4).

The simplest types of reversible thermal processes occurring in an ideal gas are analyzed on the basis of the equation

of state. The state of a gas is defined by three parameters, viz. pressure p, volume V, and temperature T. As the state of the system changes in a process, the values of p, V, and T also change, but their variation takes place in such a way that they always satisfy the equation of state. This means that the thermal processes considered here are quite slow: *at each instant of time, the system is in a new state of thermal equilibrium corresponding to the new (altered) conditions.*

If a process occurs in such a way that the temperature remains constant **(isothermal process)**, it follows from the equation of state (5.8.4) that

$$pV = \text{const.} \tag{5.8.22}$$

The process in which the pressure remains constant **(isobaric process)** is described by the relation

$$\frac{T}{V} = \text{const},$$

or

$$\frac{V}{V_0} = \frac{T}{T_0}. \tag{5.8.23}$$

Formula (5.8.23) follows from formula (5.8.4).

In accordance with (5.8.4), the processes in which the volume remains unchanged **(isochoric process)** obeys the following relations:

$$\frac{p}{T} = \text{const},$$

or

$$\frac{p}{T} = \frac{p_0}{T_0}. \tag{5.8.24}$$

Let us consider an **adiabatic process** occurring in a thermally insulated system which does work or on which work is done. We use the general thermodynamic relation (5.7.4) in which we put $dQ = 0$. The change in the state of a body in an adiabatic process is then described by the equation

$$dE + p\,dV = 0. \tag{5.8.25}$$

Let us apply this relation to the adiabatic expansion (or compression) of an ideal gas. We consider one mole of the gas. In this case, Eq. (5.8.25) can be written in the form

$$C_V \, dT + p \, dV = 0. \tag{5.8.26}$$

In this equation, we have used definition (5.8.12) of heat capacity, as well as the fact that for an ideal gas (having a fixed mass), E is a function of only T. Substituting expression (5.8.4) for p into (5.8.26) and dividing the whole equation by T, we obtain the relation

$$C_V \frac{dT}{T} + R \frac{dV}{V} = 0. \tag{5.8.27}$$

We shall consider the temperature region in which the heat capacity of the gas is constant. In this case, we can write (5.8.27) in the form

$$d\,(C_V \ln T + R \ln V) = 0. \tag{5.8.28}$$

Here, we have used the expression for the derivative of a logarithmic function and the fact that a constant quantity can be included in or taken out of a differential d. Equality (5.8.28) indicates that the expression in the parentheses is a constant, i.e.

$$C_V \ln T + R \ln V = \text{const}. \tag{5.8.29}$$

This expression can also be written in the form

$$\ln (T^{C_V} V^R) = \text{const}.$$

This means that

$$T^{C_V} V^R = \text{const}. \tag{5.8.30}$$

Finally, taking (5.8.18) into consideration and raising (5.8.30) to the power $1/C_V$, we obtain

$$TV^{\gamma - 1} = \text{const}, \tag{5.8.31}$$

where

$$\gamma = \frac{C_p}{C_V}.$$

Thus, in an adiabatic process, the temperature and volume change in such a way that the product $TV^{\gamma-1}$ remains

unchanged. Since $\gamma > 1$ in accordance with (5.8.16), and hence $\gamma - 1 > 0$, *the adiabatic expansion of a gas is accompanied by cooling, while the adiabatic compression is accompanied by heating.* This circumstance follows from the relation $dE = dA$ describing a thermally insulated system since the work of expansion is done in this case at the expense of the internal energy which is an increasing function of temperature for an ideal gas (see Eq. (5.8.7)).

Using the equation of state (5.8.4) for an ideal gas and relation (5.8.31) obtained above, we can find the relations between other variables describing an adiabatic process, viz.

$$pV^{\gamma} = \text{const} \tag{5.8.32}$$

and

$$T^{\gamma}p^{1-\gamma} = \text{const}. \tag{5.8.33}$$

Let us compare the isothermal and adiabatic expansion of an ideal gas. Equation (5.8.22) can be taken as that of the isotherm, while (5.8.32) can be taken as that of the adiabat. Suppose that the parameters of the gas at the initial moment of time are p_0 and V_0. Proceeding from this initial state, we carry out the two different processes of expansion. The results are represented in Fig. 5.14.

Earlier, in Sec. 5.7 (see Fig. 5.11), we considered the **Carnot cycle** consisting of two isotherms and two adiabats. The reader should now be in a position to understand the processes constituting this cycle.

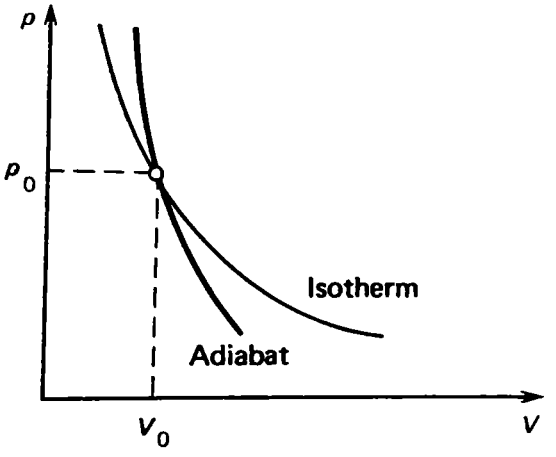


FIG. 5.14

5.9. Statistics and Thermodynamics of Radiation

The electromagnetic radiation in a cavity may be in thermal equilibrium with the material of the cavity walls. If electromagnetic radiation is treated as a collective of photons (see Sec. 4.7.2), we are dealing with an equilibrium "photon gas".

The fact that electromagnetic waves are not scattered by one another is reflected in the linear nature of the Maxwell equations. The linearity of the field equations ensures the application of the superposition principle to the field, which means that fields may be superimposed on one another without mutual influence. Consequently, a photon gas can be treated as an ideal gas.

Photons are quantum objects since equations (4.7.3), (4.7.4), and (4.7.6) for energy, angular momentum, and momentum contain Planck's constant $\hbar$. Since photons have an integral angular momentum (see Sec. 4.7.2), the equilibrium photon gas must obey Bose statistics.

Photons in a cavity are continuously absorbed by the cavity walls which, in turn, emit new photons. In the state of thermal equilibrium between the walls and the radiation, all these processes neutralize one another, and the distribution of photons over frequencies and directions remains unchanged on the average.

Suppose that the walls are made of a material consisting of different groups of atoms, each of which includes N atoms. We assume that these atoms can exist in two states only, viz. the ground state i and the excited state j. The energy difference ΔE between these states corresponds to the emission or absorption of a photon with energy $\hbar\omega$ (Fig. 5.15*a*). Thus, the photons with a given frequency ω are ascribed

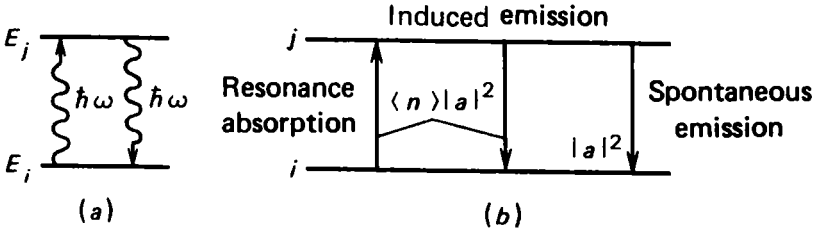


FIG. 5.15

to a certain group of atoms (the number of atoms in the group is quite large). We denote by N_{gr} and N_{ex} the mean number of atoms in groups in the ground and excited states respectively. In thermal equilibrium, these numbers must satisfy Eq. (5.6.5) which can be written in the form

$$\frac{N_{ex}}{N_{gr}} = e^{-\Delta E/(kT)} = e^{-\hbar\omega/(kT)}. \tag{5.9.1}$$

Figure 5.15*b* shows the two main processes occurring between the photons in the cavity and the atoms of the cavity walls. One of these processes (termed **direct process**) involves the absorption of a photon with frequency ω by an atom in the ground state, which thus goes over to an excited state. In contrast to this, the **reverse process** involves a transition of the atom from an excited state to the ground state, accompanied by the emission of a photon with frequency ω. Two types of emission are possible in this case, viz. **spontaneous** and **induced emission**, the probability of the latter being proportional to the number of photons with frequency ω present in the cavity (see Sec. 4.7).

In the statistical equilibrium, the rates of the direct and reverse processes must be equal. The rate of a process is determined by the number of atoms in the group, as well as by the mean number $\langle n\rangle$ of photons with frequency ω in the given state. The mean number $\langle n\rangle$ of photons appears in the expression $\langle n\rangle\,|a|^2$ for the probability of resonance absorption, and in the expression $\langle n\rangle\,|a|^2$ for the probability of induced emission; $|a|^2$ is the quantum-mechanical probability of spontaneous emission. In this case, the rate of photon absorption from a "subsystem of photons with frequency ω" is $N_{gr}\,\langle n\rangle\,|a|^2$, while the rate of photon emission to the "subsystem of ω-photons" is $N_{ex}\,(\langle n\rangle + 1)\,|a|^2$, where unity in the expression within the parentheses describes spontaneous emission, and $\langle n\rangle$ corresponds to induced emission. Equating the rates of the two processes, we obtain

$$N_{gr}\,\langle n\rangle = N_{ex}\,(\langle n\rangle + 1). \tag{5.9.2}$$

Comparing (5.9.2) and (5.9.1), we obtain

$$\frac{\langle n\rangle}{\langle n\rangle + 1} = e^{-\hbar\omega/(kT)},$$

or

$$\langle n \rangle = \frac{1}{e^{\hbar\omega/(kT)} - 1}. \tag{5.9.3}$$

It should be observed that (5.9.3) is the analytic expression for Bose distribution as applied to a photon gas.

Although the derivation of (5.9.3) is based on a certain model, this distribution is directly connected with reality. It is essentially of quantum origin and describes the distribution of photons at temperature T over different quantum states with energies $\varepsilon = \hbar\omega$, where ω are the natural radiation frequencies in a given volume V of the cavity. Assuming this volume to be quite large, we can go over from a discrete distribution of the natural radiation frequencies to a continuous distribution.

Suppose that the volume V of a cavity in which free oscillations of a field are established (see Sec. 3.8.4) is not small. This means that the waves freely propagate under the condition that $\lambda \ll L$, where L represents the linear dimensions of the cavity. We introduce the six-dimensional "phase state space". The elementary phase cell is defined as

$$\Delta p_x \Delta p_y \Delta p_z \Delta x \Delta y \Delta z \sim \hbar^3,$$

where we have used the uncertainty relations (4.3.2), and p indicates the momentum. For a photon, the modulus of whose momentum lies between p and $p + \mathrm{d}p$, the number of quantum states is

$$2\,\frac{4\pi p^2\,\mathrm{d}pV}{\hbar^3},$$

where the factor 2 corresponds to two independent directions of photon polarization (see Sec. 4.9). Since the momentum of a photon is $p = \hbar\omega/c$, the number of quantum states with frequencies between ω and $\omega + \mathrm{d}\omega$ is

$$\frac{V\omega^2\,\mathrm{d}\omega}{c^3}. \tag{5.9.4}$$

Here, we have neglected the numerical factors. Multiplying distribution (5.9.3) by (5.9.4), we obtain the following ex-

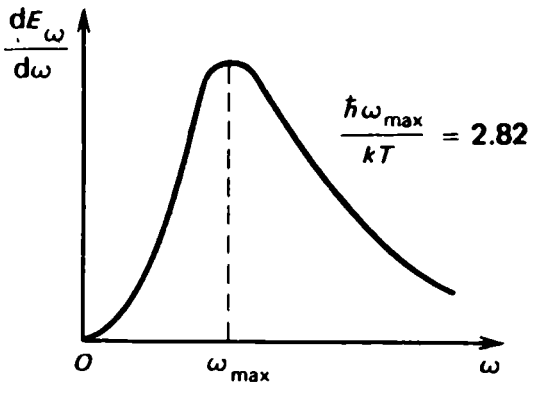


FIG. 5.16

pression for the number of photons in a given frequency interval:

$$dN_\omega = \frac{V}{c^3} \frac{\omega^2 \, d\omega}{e^{\hbar\omega/(kT)} - 1}. \tag{5.9.5}$$

Multiplying this relation by $\hbar\omega$, we obtain the radiant energy corresponding to this interval of the spectrum:

$$dE_\omega = \frac{V\hbar}{c^3} \frac{\omega^3 \, d\omega}{e^{\hbar\omega/(kT)} - 1}. \tag{5.9.6}$$

This formula, called **Planck's formula**, gives the spectral distribution for the energy of thermal radiation (see the curve represented in Fig. 5.16).

Figure 5.16 shows the position of the maximum of $dE_\omega/d\omega$ (spectral energy distribution density of thermal radiation over frequencies). If we consider the energy distribution of thermal radiation over wavelengths ($\lambda = 2\pi c/\omega$), the quantity $dE_\lambda/d\lambda$ attains its maximum value when the following condition is satisfied:

$$2\pi \left(\frac{\hbar c}{k}\right) \frac{1}{T\lambda_{max}} = 4.96, \text{ [2]}$$

or

$$T\lambda_{max} = \text{const} = \frac{2\pi\hbar c}{4.96k}. \tag{5.9.7}$$

---

[2] This numerical result is obtained for $\lambda = \lambda_{max}$ by solving the equation $dE_\lambda/d\lambda = 0$ (mathematical extreme-value problem).

As the temperature of a body emitting radiation is increased, the maximum of the radiant energy shifts towards shorter wavelengths in the spectrum.

In order to estimate the significance of formula (5.9.7), let us calculate by way of an example the temperature of the outer layers of the Sun. It is well known that the maximum of the solar spectrum lies in the yellow-green region of the spectrum, i.e. $\lambda_{max} \simeq 5.5 \times 10^{-7}$ m. In this case, we obtain from (5.9.7)

$$T \simeq \frac{6.28 \times 1.05 \times 10^{-34} \times 3 \times 10^{8}}{4.96 \times 1.38 \times 10^{-23} \times 5.5 \times 10^{-7}} \text{ K} \simeq 5300 \text{ K}.$$

Integrating (5.9.5) over all frequencies (from 0 to ∞), we obtain

$$dN_{\omega} \rightarrow N,$$

where

$$N = 0.244 \left(\frac{kT}{\hbar c}\right)^3 V \tag{5.9.8}$$

is the total number of photons in thermal radiation. The integration can be carried out graphically by plotting function (5.9.5) and computing the area under this curve. It follows from (5.9.8) that

$$VT^3 = \text{const.} \tag{5.9.9}$$

This is a thermodynamic relation.

The second thermodynamic relation for thermal electromagnetic radiation is obtained by integrating (5.9.6) in the same unlimited frequency range. This gives

$$dE_{\omega} \rightarrow E,$$

where

$$E = 0.658 \frac{(kT)^4}{(\hbar c)^3} V \tag{5.9.10}$$

is the total energy of thermal radiation. Since the pressure

$$p \sim \frac{E}{V} \tag{5.9.11}$$

(see the beginning of Sec. 5.8.3), it follows from (5.9.10) and (5.9.11) that

$$\frac{p}{T^4} = \text{const.} \tag{5.9.12}$$

Solving (5.9.9) and (5.9.12) simultaneously and eliminating T from these equations, we obtain

$$pV^{4/3} = \text{const.} \tag{5.9.13}$$

It should be observed that the volume, pressure, and temperature of a thermal photon gas under adiabatic expansion (or compression) are connected through relations (5.9.9), (5.9.12), and (5.9.13). These relations are found to be useful for a large number of cases, for example, in studying the evolution of stars since the light pressure in a star plays a significant role.

Let us determine the **surface flux density of thermal radiation (radiation intensity)**, i.e. the ratio of the energy emitted per unit time to the area of the emitting surface. For this purpose, we proceed from the relation

$$j \sim \frac{E}{V} c, \tag{5.9.14}$$

where E/V is the volumetric energy density of the field, c is the velocity of light (see Sec. 3.7.2), and the dimensions of j are $[j] = 1\ \text{J/(m}^2 \cdot \text{s)}$. Obviously, to determine the value of j, it is sufficient to multiply the radiation energy density in (5.9.10) by the constant c. This gives

$$j \sim \left(\frac{k^4}{\hbar^3 c^2} \right) T^4. \tag{5.9.15}$$

Exact calculations lead to the expression

$$j = \sigma T^4, \tag{5.9.16}$$

where $\sigma = 5.67 \times 10^{-8}\ \text{W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$ is called the **Stefan-Boltzmann constant.** Thus, the intensity of the thermal radiation emitted by a square metre of the emitting surface is defined by formula (5.9.16).

Note that formula (5.9.16) can also be derived purely thermodynamically, but the numerical value of σ cannot be obtained by this method. It was mentioned above that the

statistical approach provides not only the relations between quantities but also the necessary numerical constants. In thermodynamics, the constant σ can be determined only experimentally.

5.10. Crystals

5.10.1. Crystal Lattice. In the state of complete thermodynamic equilibrium at $T = 0$ K, the energy of a system has its minimum value, and only a few selected (ordered) positions from among the possible mutual arrangements of particles should be observed. In other words, *all solids must be crystalline at low temperatures* (see Sec. 6.14 for the only exception to this rule).

By a **crystal (crystalline solid)**, we mean above all the periodic nature of its microscopic structure. Each atom in a crystal is surrounded by atoms arranged in a definite manner relative to one another. If this configuration of atoms has the lowest possible energy, it must obviously be repeated in other parts of the body as well.

The simplest configuration of atoms (generally speaking, atoms of different types) which is repeated periodically over a body in all three dimensions forms the **unit cell of the crystal lattice** (Fig. 5.17).

The crystal lattice is said to possess **translational symmetry** with periods **a**, **b**, and **c** along three directions (see Fig. 5.17). Naturally, the physical quantities like the charge

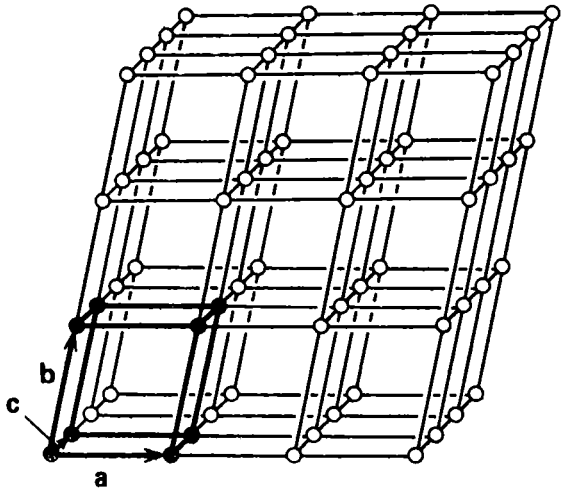


FIG. 5.17

density of atomic electrons have the same periodicity as the lattice itself.

To find out the number of crystallographic systems that may exist in nature, we must consider the possible types of symmetry in space lattices. We shall not go into details of this interesting topic and confine ourselves to the remark that the number of types of lattice symmetry is limited.

This can be verified by using the following example. Let us consider the problem of filling a plane with identical figures in such a way that the resulting pattern is free of voids between the figures. Obviously, the figures cannot have an arbitrary geometry. The sum of the angles of adjacent figures about a common vertex must be equal to 360°, and the angles themselves must be integral fractions of 360°, i.e. 60, 90, and 120° in the simplest case. As a result, symmetric patterns may be formed by rectangles, triangles, squares, and hexagons (regular, of course). This example is applicable to real crystals as well since a plane lattice is a projection of the space lattice.

5.10.2. Types of Lattice Bonds. Different types of bonds **(ionic, covalent, metallic, van der Waals)** exist in atomic lattices (see Sec. 4.10). A typical example of a crystal with an ionic bond is the NaCl crystal which is formed as a result of a redistribution of the outer atomic electrons between sodium and chlorine atoms. The valence electron of Na goes over to the outer shell of a chlorine atom and completes it. As a result, the NaCl crystal is formed not by sodium and chlorine atoms but by Na^+ and Cl^- ions. The force of electrostatic attraction between these ions ensures the bond between sodium and chlorine in the NaCl crystal.

The nature of covalent bonds in a crystal can be explained by considering two approaching atoms. When the atoms are separated by a large distance, each electron experiences the action of the field of its own nucleus only. When the separation between the atoms is reduced to such an extent that their electron probability amplitudes $\psi(r)$ (see Fig. 4.13*a*) begin to overlap, a common electron state of two electrons with antiparallel spins in the field of both nuclei is formed. The energy of such an electron pair is lower than their total energy in the isolated states (see Sec. 4.10). Such a bond is called a covalent bond.

Diamond is a typical crystal with covalent bonds. The crystal lattice of diamond consists of carbon atoms having four valence electrons each. Each carbon atom lies at the centre of a tetrahedron whose vertices are formed by four nearest atoms. This means that each atom in the diamond lattice is bound in four directions through covalent bonds. The high energy of a covalent bond and its directionality make a covalent lattice quite rigid.

A metallic bond is of a quite different nature. This bond does not appear between nearest atoms but is rather a property of the entire crystal. The outer (valence) electrons in a metal crystal are fully collectivized and form what is known as an **electron gas**. Positively charged ions occupy the lattice sites in the crystal. Metallic bonds are not strictly directional in nature.

In molecular crystals, the molecules retain their individuality and are connected with one another through weak van der Waals forces. Such crystals differ sharply from crystals like diamond which is just a huge molecule. The van der Waals type of crystal also includes monatomic crystals like solid argon.

The binding energy of particles in a van der Waals crystal is an order of magnitude or two lower than in crystals of any other type. Hence van der Waals crystals are characterized by very low melting points (usually from —150 °C to +350 °C). The high porosity of their structure is responsible for a high compressibility of such crystals.

Several different types of bonds are simultaneously observed in crystals. For example, in the case of graphite which has a complex structure, covalent bonds are formed (separation 0.142 nm) in planes constituted by hexagonal lattices, while van der Waals forces act between layers (separation 0.34 nm). Such a strange combination of different types of bonds in the graphite lattice results in a remarkable combination of its properties: graphite has a low hardness and at the same time one of the highest melting points (about 4000 °C).

5.10.3. Mechanical Properties of Crystals. The energy U of interaction between atoms as a function of the interatomic separation r is different in different lattices and also depends on the type of the atoms forming a lattice. Leaving

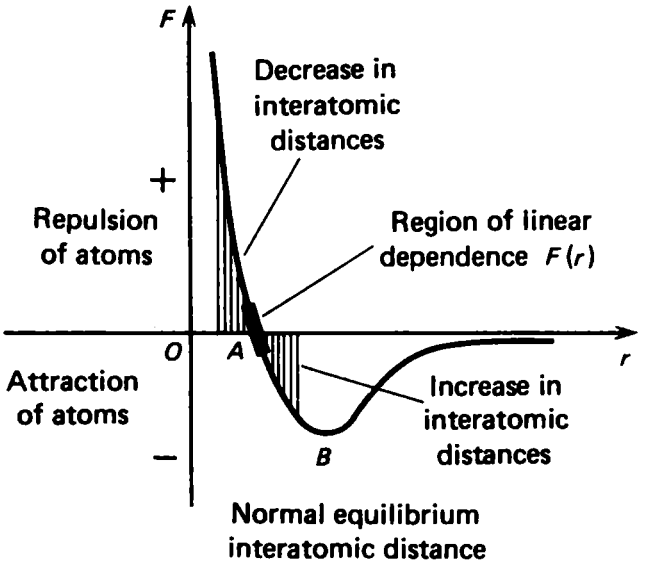


FIG. 5.18

aside the details, we can plot a characteristic dependence $U(r)$ for any type of atomic interaction. This dependence has the form shown in Figs. 3.28 (see the variation of U_{eff}) and 6.5.

For our purpose, it is sufficient to analyze the curve $F(r)$, where

$$F = -\frac{dU}{dr}$$

is the force of atomic interaction. All the remarks made in connection with the function $U(r)$ are naturally valid for the function $F(r)$ which is the derivative of U. The shape of the function $F(r)$ is shown in Fig. 5.18.

We shall consider the macroscopic representation of the function $F(r)$ as the dependence of the elastic force $|F_{\text{el}}|$ on the deformation Δl (the region AB in Fig. 5.18 is mapped by the segment $A'B'$ in Fig. 5.19).

Let us formally derive a simple and convenient approximate expression for the function $F_{\text{el}}(\Delta l)$. We proceed from the general expression for the derivative of the function $F_{\text{el}}(\Delta l)$ with respect to Δl at the point $\Delta l = 0$:

$$F'_{\Delta l}(0) = \lim_{\Delta l \to 0} \frac{F(\Delta l) - F(0)}{\Delta l}. \qquad (5.10.1)$$

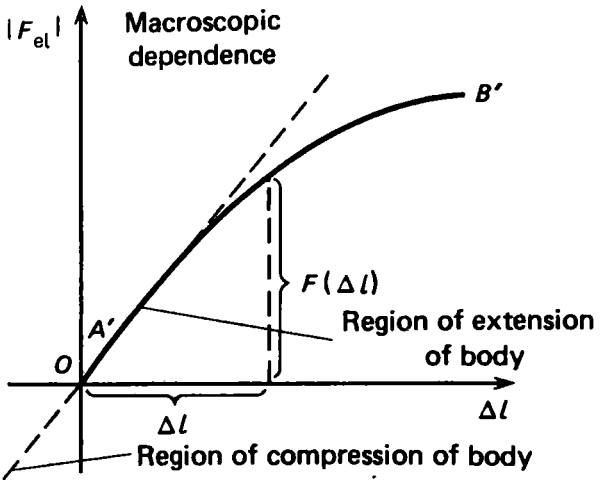


FIG. 5.19

This expression can be written in the form

$$F(\Delta l) \approx F(0) + F'_{\Delta l}(0)\,\Delta l. \tag{5.10.2}$$

This is the approximate expression for the function $F(\Delta l)$ at the point Δl in terms of its value and the value of its derivative at zero. Since the derivative $F'_{\Delta l}$ varies in the interval Δl, the exact expression for $F(\Delta l)$ at the point Δl has the form

$$F(\Delta l) = F(0) + \int_0^{\Delta l} F'_{\Delta l}(\Delta l)\,\mathrm{d}\,(\Delta l). \tag{5.10.3}$$

Taking the derivative of (5.10.2),

$$F'_{\Delta l}(\Delta l) \approx F'_{\Delta l}(0) + F''_{\Delta l}(0)\,\Delta l,$$

and substituting it into (5.10.3), we obtain

$$F(\Delta l) \approx F(0) + \int_0^{\Delta l} [F'_{\Delta l}(0) + F''_{\Delta l}(0)\,\Delta l]\,\mathrm{d}\,(\Delta l)$$

$$= F(0) + F'_{\Delta l}(0)\,\Delta l + F''_{\Delta l}(0)\,\frac{(\Delta l)^2}{2}. \tag{5.10.4}$$

Carrying out similar operations for refining the approximate expression (5.10.2), we obtain from (5.10.4)

$$F(\Delta l) \approx F(0) + F'_{\Delta l}(0)\,\Delta l + F''_{\Delta l}(0)\,(\Delta l)^2 + F'''_{\Delta l}(0)\,(\Delta l)^3. \tag{5.10.5}$$

In this formula, we have omitted the numerical coefficients of the second and third derivatives. It is obvious from the derivation that formula (5.10.5) is applicable not only for $F(\Delta l)$ but for any smooth continuous function. This formula can be extended indefinitely through the addition of subsequent terms. Such a mathematical operation is called **expansion of a function into a series.** This is a remarkable achievement of mathematicians which serves as an effective tool in the hands of physicists.

Let us illustrate this statement. Let us impose the conditions of a specific physical problem on the general mathematical formula (5.10.5). These conditions can be formulated as follows: (i) in the state of equilibrium, when there is no deformation (i.e. when $\Delta l=0$), elastic forces do not emerge, i.e. $F(0)=0$; (ii) since bodies exhibit symmetry upon extension and compression (not too far from the equilibrium state), the function $F(\Delta l)$ must be odd, i.e. $F(\Delta l) = -F(-\Delta l)$, which means that all the terms with even powers are absent in the expansion with respect to small deformations. Consequently, we get

$$F(\Delta l) \approx F'_{\Delta l}(0)\,\Delta l - F'''_{\Delta l}(0)\,(\Delta l)^3,$$

where

$$F'_{\Delta l}(0) = \text{const} \equiv \alpha,$$

$$F'''_{\Delta l}(0) = \text{const} \equiv \beta.$$

For small deformations,

$$\alpha\,\Delta l \gg \beta\,(\Delta l)^3,$$

and the dependence

$$F_{\text{el}}(\Delta l) \approx \alpha\,\Delta l \qquad (5.10.6)$$

is linear **(Hooke's law).** For large deformations,

$$\alpha\,\Delta l \gtrsim \beta\,(\Delta l)^3$$

and

$$F_{\text{el}}(\Delta l) \approx \alpha\,\Delta l - \beta\,(\Delta l)^3, \qquad (5.10.7)$$

which means that the dependence is no longer linear. In (5.10.6), as well as (5.10.7), the deformation is assumed to

be elastic, i.e. the body is assumed to retrieve its shape and size after the mechanical loads have been removed.

From the thermodynamic point of view, the deformed state of bodies is a nonequilibrium state. Hence internal stresses appear in a deformed body and tend to revert it to the state of thermal equilibrium. This means that any mentally isolated region at any position along the body is subjected to such a force from all the surrounding parts that the resultant of all the forces is zero.

It should be observed that the internal stresses in a deformed body are caused by the atomic interactions. Naturally, there is no internal strain in a body in the undeformed state.

Let us determine the limiting internal stresses in *ideal* crystalline solids, i.e. in **defect-free single crystals.** We shall consider the limiting mechanical strength of ideal ionic crystals. The internal stresses p characterizing the rupture of a body may be presented in the form (see Sec. 5.8.3)

$$p \sim \frac{E}{V} \sim \langle U \rangle \frac{N_1}{V},$$

where (N_1/V) is the number density of ions in the body. The mean interaction energy for a pair of singly charged ions with opposite charges is $\langle U \rangle \sim e^2/r$, which is of the order of 10^{-18} J. In order to evaluate $\langle U \rangle$ in the above formula, we first substitute the values of the quantities expressed in the CGS units and then convert them into the SI units. Since $(N_1/V) \sim 10^{28}$ ions/m^3, we obtain

$$p \sim 10^{10} \text{ J/m}^3.$$

The values of the mechanical strength obtained in this way exceed the real strength of materials by a factor of 10^2-10^3. The reason behind this discrepancy between the theoretical estimates and the results of actual measurements will be described at a later stage. For the present, we simply note that the mechanical strength of defect-free crystals in the form of whiskers is close to the theoretical value.

Subjecting an ideal crystal to uniform deformation (under which each minute part of the crystal is deformed in the same manner) at low temperatures, the dependence of the

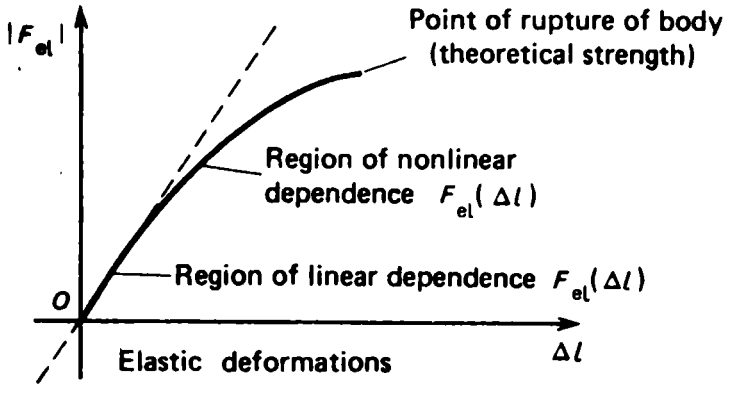


FIG. 5.20

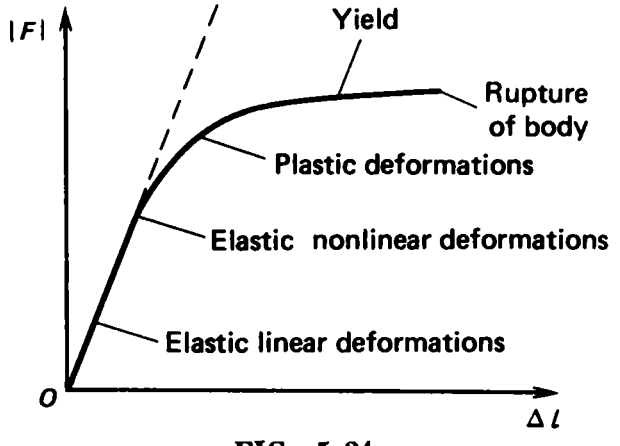


FIG. 5.21

internal elastic forces F_{el} on the deformation Δl must have the form shown in Fig. 5.20. It should be noted that deformations in an ideal crystal are reversible, i.e. purely elastic.

Let us consider *real* crystals. In this case, we shall carry out essentially a qualitative survey. In addition to elastic (reversible) deformations, irreversible, or **plastic deformations**, also emerge in real crystals (for considerable loads). Such deformations produce residual changes in the crystal after the removal of the load.

While analyzing the mechanical properties of real crystals, we proceed with the form of the dependence F (Δl) for them. Figure 5.21 shows the qualitative behaviour of this dependence. The curve contains regions of elastic (linear and nonlinear) deformations, as well as plasticity and yield regions. A comparison of the curve for real solids (see

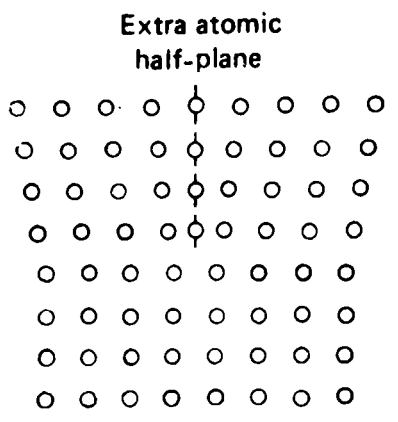


FIG. 5.22

Fig. 5.21) with the corresponding curve for ideal crystals (see Fig. 5.20) should be carried out bearing in mind that the two curves are plotted on entirely different scales. Indeed, it was mentioned above that the maximum value of the quantity $|F|$ for ideal crystals is hundreds or even thousands of times larger than the corresponding value for real solids. This is so because, on the one hand, we are dealing with a strict order in the arrangement of atoms in a lattice, while, on the other hand, the regularity in the packing of atoms in a crystal is violated, i.e. **lattice defects** are present in real crystals.

Defects in crystalline solids are classified as follows: (i) a group of **point defects** which include **vacancies**, or unoccupied lattice sites, and **interstitial atoms** (belonging to the basic matter comprising the crystal or to a "foreign" impurity), and (ii) **linear defects** which include **dislocations.** Dislocations are lattice defects caused by the presence in a crystal of an extra half-plane wedged between two geometrically regular atomic planes (Fig. 5.22).

The mechanical properties of real solids are mainly determined by the behaviour of dislocations in them. The loss of elastic properties in real samples is due to the appearance of dislocations, their movement and interaction with other dislocations or defects of other types.

To illustrate the connection between the macroscopic behaviour of a mechanically loaded sample and the movement of dislocations in it, let us consider the **strain hardening of metals** upon considerable plastic deformations. To begin with, we observe that any type of uniform deformation of crystalline solids results in a **uniform compression** (or **extension**) and **shear**. A body under a uniform compression retains its shape, but its volume decreases. The situation is reverse in the case of shear: the volume of the body remains the same, but its shape changes. The shear component is responsible for the emergence of slip dislocation

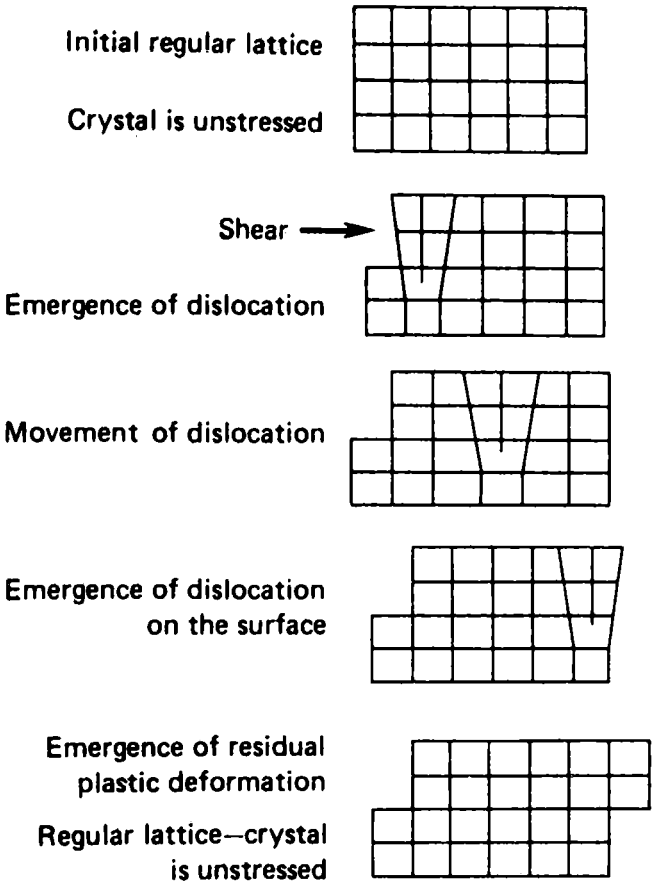


FIG. 5.23

(Fig. 5.23) at comparatively low stresses (the slipping of one part of the crystal as a whole relative to the other part is possible at very high stresses). If the sample being tested has a fairly ordered structure, a dislocation which has started moving is practically not arrested, and the plastic deformations grow. But the growth of deformation results in the appearance of new dislocations which hinder the free slipping of other dislocations. The arrest of a dislocation corresponds to a termination of plastic deformation, and hence to an increase in the elastic limit of the body. This phenomenon is called strain hardening (segments *0-1-2* and *2-3* in Fig. 5.24 correspond to loading and unloading respectively, segment *0-3* describes the residual deformation, while *3-4* is the region of elastic deformation).

There are several methods of direct experimental observation of dislocations. The most direct method involves

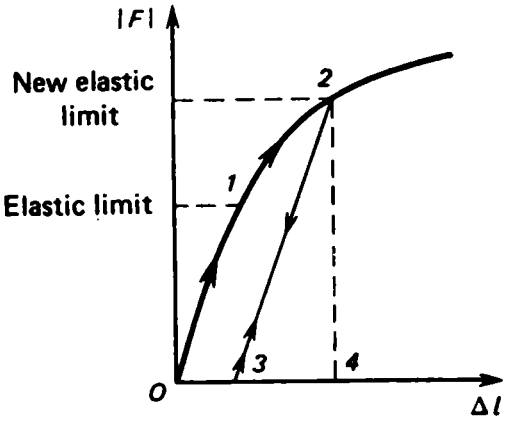


FIG. 5.24

the use of an electron microscope which allows us to visually observe the arrangement of atomic planes including the "extra" atomic half-planes, viz. dislocations.

Let us consider in brief the **yield of solids**. Under limiting (high) mechanical loading, a considerable increase in the volume of solids takes place owing to the formation of "vacant sites" in the lattice, i.e. owing to vacancies. For high internal stresses, the vacancies actively diffuse through the "damaged" crystal in the direction of the applied external mechanical forces. A situation resembling the flow of a highly viscous liquid is observed, but the rate of the flow is extremely small. In experiments, this is manifested in the form of a continuous increase in the deformation of the body without any increase in the applied load.

This mechanism of yield of solids is realized at high temperatures. At low temperatures, plastic flow of dislocations apparently takes place.

5.10.4. Electron Energy Spectra of Crystals. The outer (valence) electrons of the atoms in a crystal may be collectivized in different ways as a result of the interaction of the outer electron shells of these atoms. The existence of such an interaction indicates that we cannot speak of the energy levels of individual atoms in the solids, and consider only the levels of the aggregate of electron shells of all the atoms constituting the solid.

A lattice consisting of N cells contains N identical and uniformly arranged atoms. An electron e^- with the same

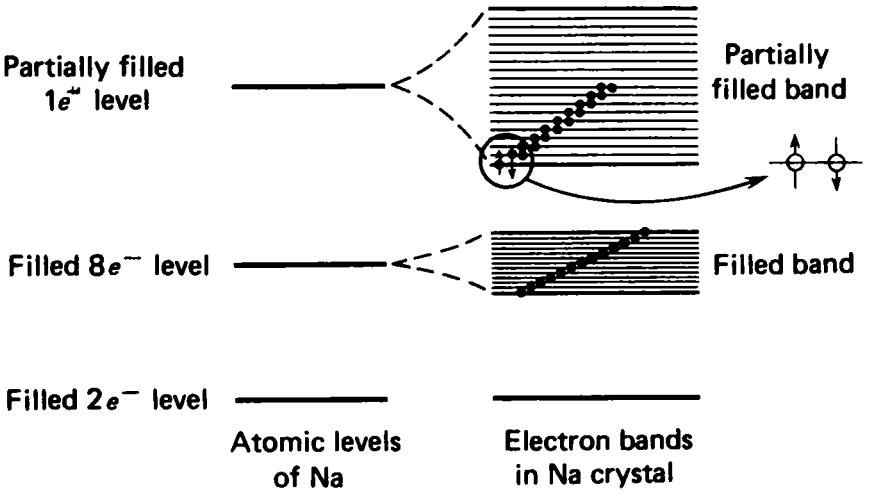


FIG. 5.25

energy can belong to any one of these atoms (the ψ functions of these atoms overlap). Hence a simple atomic level acquires an additional N-fold degeneracy in the lattice. This "transposition" degeneracy is removed as a result of atomic interaction in the lattice, and the atomic level is transformed into an energy band consisting of N closely spaced levels (Fig. 5.25).

The extent to which the electron levels are populated in any energy band depends on the population density of the corresponding atomic level. If all the quantum states with energy E_n in an atom are filled by electrons, the band corresponding to this level is also filled.

Vacant levels give rise to vacant bands, while partially filled levels create partially filled bands.[3] Solids with partially filled bands are called **metals**, and those with an energy spectrum of electron states consisting of filled or empty bands are known as **insulators** (Fig. 5.26).

Owing to the presence of a partially filled band in metals, the thermal motion of electrons may cause their redistribution over different states (electrons may go over to the neighbouring vacant levels in the band). On the contrary, electrons of a filled band in an insulator cannot take part in the thermal motion (on account of the absence of vacant levels in the band).

[3] Of course, there are exceptions to this rule since the bands may overlap.

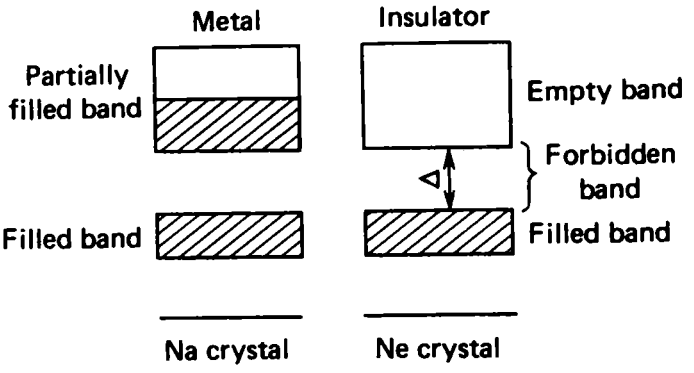


FIG. 5.26

Unlike metals, the electron spectrum in an insulator is characterized by the presence of an **energy gap (forbidden band)**. Hence the properties of insulators are characterized by two bands, viz. a filled band and an empty band, and depend considerably on the gap separating the two bands. If the width of the forbidden band is not very large, the term **semiconductor** is used instead of insulator.

The energy gap or the forbidden band in insulators is of the order of several electron volts (see Fig. 5.26). This value corresponds to a temperature of tens of thousand degrees. At normal temperatures for insulating crystals, when $kT \ll \Delta$, Δ being the band gap, the number density of free electrons arriving in the "empty" band from the filled band as a result of thermal motion is extremely low and depends exponentially on temperature ($\sim \exp[-\Delta/(kT)]$). The band gap Δ in semiconductors is $\lesssim 1$ eV.

Thus, the band structure of the energy spectrum of electron states in crystals explains the basic fact of the existence of three main types of crystalline solids in nature, viz. metals, insulators, and semiconductors. This terminology is connected with the electrical conductivity of bodies (see Chap. 7).

5.10.5. Lattice Heat Capacity. The main contribution to the energy of thermal motion is made by the lattice, the electrons playing a very insignificant role.[4] This is due to the

[4] With the exception of metals at temperatures close to absolute zero (see Sec. 5.10.6).

large difference in the numbers of degrees of freedom of the lattice and the electrons. In an insulator at normal temperatures, all electrons are confined to filled bands and can neither change their position nor participate in the thermal motion. Electrons in metals can change their energy states. However, in spite of the fact that the number of electrons participating in the thermal motion increases with temperature, it is very small in comparison with the number of nuclei in the lattice (see Sec. 5.10.6 for quantitative estimates of these parameters). Hence the thermodynamic quantities obtained below completely describe the thermal properties of insulating crystals and only the lattice component of the heat capacity of metals.

The thermal motion of nuclei in a lattice can be described by their vibrational motion about certain equilibrium positions (lattice sites). This motion does not cease even when the temperature drops to absolute zero. The vibrations of the nuclei in a lattice at absolute zero are sometimes called zero-point vibrations.

Let us determine the temperature interval in which the vibrations of lattice nuclei must be treated quantum mechanically. According to (5.8.10), we have

$$T_{\mathrm{d}} \sim \frac{\hbar^2}{m_p k}\left(\frac{N_p}{V}\right)^{2/3}, \tag{5.10.8}$$

where m_p is the proton mass. Substituting the numerical values, we get

$$T_{\mathrm{d}} \sim \frac{(10^{-34})^2}{10^{-27}\times 10^{-23}}(10^{28})^{2/3}\ \mathrm{K} = 10\ \mathrm{K}.$$

Thus, lattice vibrations must be treated quantum mechanically at temperatures $T \lesssim 10$ K. At higher temperatures, lattice may be analyzed by classical methods.

Weakly excited lattice states, i.e. states at low temperatures, are characterized by the appearance of collective elastic vibrations of low frequency ω in the lattice. These vibrations are nothing but ordinary acoustic waves which are of macroscopic nature. Hence their wavelength λ must be much longer than the lattice constant a (see Fig. 5.17). In turn, the condition $\lambda \gg a$ means that

$$\omega \ll \frac{u}{a}, \tag{5.10.9}$$

where u is the velocity of sound. While deriving this equation, we have used the formula $\lambda \sim u/\omega$. Thus collective vibrations of nuclei in the lattice can be considered acoustic waves if the temperature satisfies the condition

$$\frac{kT}{\hbar} \ll \frac{u}{a},$$

or

$$T \ll \frac{\hbar u}{ka}. \tag{5.10.10}$$

In order to determine the temperature region corresponding to (5.10.10), we put $u \sim 10^3$ m/s and $a \sim 10^{-10}$ m in this inequality. This gives

$$T \ll \frac{10^{-34} \times 10^3}{10^{-23} \times 10^{-10}} \text{ K} = 10^2 \text{ K}. \tag{5.10.11}$$

This estimate is in accord with (5.10.8).

Let us now introduce the concept of sound quanta, or **phonons**. The idea about phonons appears quite naturally in this analysis. In order to justify the introduction of this new concept, we observe that while the nuclei perform individual intramolecular vibrations in the diatomic molecules of an ideal gas (see Sec. 4.10), the vibrations of nuclei in a lattice are of collective type. *The concept of phonons arises as a means of quantum description of the collective vibrations of nuclei in a lattice.*

Phonons are treated as **quasi-particles** since, unlike, say, photons and electrons, they cannot exist outside the crystal. Collective lattice vibrations are equivalent to a "phonon gas" in which each phonon propagates freely within the volume occupied by a body. The energy of a phonon is

$$\varepsilon = \hbar\omega, \tag{5.10.12}$$

where ω is the frequency of an elastic wave; the phonon momentum satisfies the following relation (for small values of p)[5]:

$$p \approx \frac{\varepsilon}{u}. \tag{5.10.13}$$

---

[5] As the momentum increases, the dependence $\varepsilon(p)$ becomes nonlinear, and the dispersion relation is determined entirely by the interaction between the lattice atoms. A general expression for this relation cannot be obtained in this case.

Relation (5.10.12) is analogous to the corresponding relation (4.7.4) for light quanta (photons). However, as was mentioned above, relation (5.10.13) is not identical to (4.7.6). This is so because, unlike the photon momentum, the momentum of a phonon is not a real momentum but something that can be termed as **quasi-momentum**. The essence of this statement lies in the following. The very existence of the momentum of a particle as a physical quantity is a consequence of the homogeneity of space with respect to this particle (see Secs. 1.3.2, 1.3.3, and 2.6.3). The space in a lattice for, say, an electron has a "cell symmetry" which is repeated periodically during the motion of the electron. Consequently, the momentum of a free electron in a crystal, unlike the momentum of a free electron outside a crystal, is defined ambiguously, up to a certain constant vector, called the **inverse lattice vector b**, which is connected with the lattice symmetry.

A similar analysis can be carried out for phonons as well. Consequently, the form of the **dispersion relation** $\varepsilon(\mathbf{p})$, i.e. the dependence of the energy ε of an elementary excitation on the quasi-momentum $\mathbf{p}$ in a lattice, is not only different for different crystals but also depends on the choice of the direction of quasi-momentum of phonons in the same crystal. For example, if the unit cell of a lattice consists of atoms of the same sort, it will contain three branches of $\varepsilon(\mathbf{p})$ for **acoustic vibrations**: the branch corresponding to longitudinal sound in the crystal is associated with its compressive or tensile strain, and two branches corresponding to transverse sound are shear waves.

The unit cells of complex lattices may contain two or more types of atoms. In this case, not only the displacements of neighbouring atoms in one direction should be in phase so that the unit cell is displaced as a whole but also different types of neighbouring atoms in a unit cell may undergo displacements in different directions in such a way that the centre of gravity of the unit cell remains fixed. This type of lattice vibrations form the **optical branch**[6] of $\varepsilon(\mathbf{p})$. This term arises because an oscillating electric dipole moment appears as a result of vibrations of this

[6] There are $3(k-1)$ such branches, where k is the number of atoms in a unit cell.

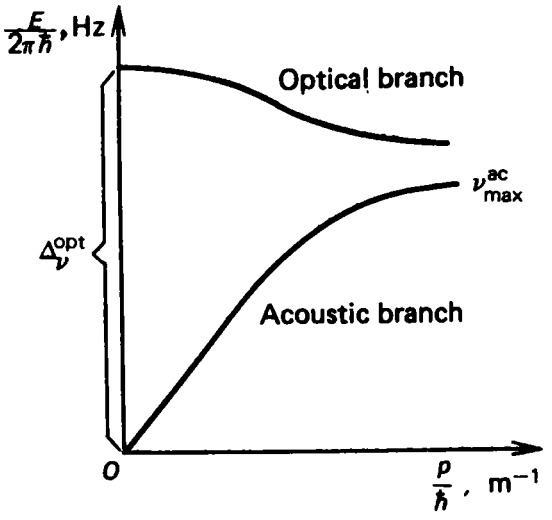


FIG. 5.27

type in ionic crystals (ions of opposite signs are displaced relative to one another periodically in antiphase). The dipole moment of the lattice actively interacts with light (usually in the infrared region of the spectrum), hence the term "optical branch".

The dispersion relations for the acoustic and optical branches have quite different forms (Fig. 5.27). In the acoustic branches, $\varepsilon \to 0$ as $p \to 0$ since in the limit these vibrations are ordinary macroscopic elastic vibrations of the medium. In the optical branches, however, ε tends to a certain finite value ε_0 as p tends to zero. Relation (5.10.13) is not applicable for optical phonons even for low values of p.

The maximum acoustic frequencies of lattice vibrations are $\nu_{max}^{ac} \sim 10^{13}$ Hz (see Fig. 5.27). This estimate can be obtained from the relation $\nu = u/\lambda_{min}$, where u is the velocity of sound, and λ_{min} is the limiting (minimum possible) length of the acoustic waves, which is equal to the lattice constant in the order of magnitude.

The energy of optical lattice vibrations is of the same order of magnitude as the vibrational energy of nuclei in a diatomic molecule, i.e. $\varepsilon_{opt} \sim 0.1$ eV (see Sec. 4.10). Accordingly, the frequency gap for exciting optical lattice vibrations is $\Delta_\nu \simeq 2 \times 10^{13}$ Hz (infrared region), see Fig. 5.27.

In lattices at low temperatures, only acoustic vibrations are excited. The fraction of the optical branches excited in the vibrational spectrum of the lattice is exponentially small.

The phonon dispersion relation is the basic dynamic characteristic of the crystal lattice. If the forces of atomic interaction are known, the dispersion relation for the lattice can be determined. However, this requires cumbersome calculations and is possible only with the help of a computer. Even in this case, the problem must be considerably simplified and the computations are only of model type.

This explains the considerable significance attached to the experimental methods of reconstructing the phonon spectra of crystals. Experiments on the scattering of slow neutrons were found to be remarkable. These experiments became possible only after the advent of nuclear reactors which are powerful sources of thermal neutrons. High-intensity beams of thermal neutrons with energy $\sim 10^{-2}$ eV "rock" the lattice atoms while passing through the crystal, i.e. excite acoustic waves in the crystal lattice (this is possible since the neutrons and the nuclei in the lattice have comparable masses and energies). In other words, inelastically scattered neutrons "generate" phonons. Using the energy and momentum conservation laws, we can determine the energy $\hbar\omega$ and momentum p_{ph} of a phonon:

$$\left.\begin{aligned} \hbar\omega &= E_n - E'_n, \\ \mathbf{p}_{\text{ph}} + 2\pi\hbar\mathbf{b} &= \mathbf{p}_n - \mathbf{p}'_n. \end{aligned}\right\}$$

Here, E_n and $\mathbf{p}_n$ are the energy and momentum of a neutron in a monochromatic beam (before scattering), E'_n and $\mathbf{p}'_n$ are the energy and momentum of a neutron scattered in the crystal, and $\mathbf{b}$ is the inverse lattice vector. The dispersion relation for the vibrations of the crystal lattice is obtained by registering the inelastically scattered neutrons.

It is well known (see Sec. 5.9) that photons in a cavity form a gas of noninteracting particles. Hence their thermal equilibrium can be attained only in a certain amount of matter. Lattice phonons do not interact only in the first approximation. In the next approximation, we must take into consideration all types of processes of elastic and inelastic scattering of phonons by phonons. It is these pro-

cesses that are responsible for the establishment of statistical equilibrium in the thermal motion of the lattice.

Such a quantum behaviour of the elastic waves in a crystal is in accord with the fact that, unlike electromagnetic waves, elastic waves may be scattered by one another (this is especially noticeable in high-intensity waves). This is due to the fact that the equations of hydrodynamics are nonlinear (see Chap. 6), while the quantitative relations describing elastic waves are derived from equations of hydrodynamics only. [7]

The number of identical phonons that can be excited in a lattice is not limited in any way. This means that *the phonon gas obeys Bose statistics.* Taking this circumstance into consideration, as well as the above-mentioned analogy between phonons and photons, we can use the method described in Sec. 5.9 for determining the heat capacity of the lattice.

Since the heat capacity at constant volume $C_V = \mathrm{d}E/\mathrm{d}T$ (see Eq. (5.8.12)), we must use expression (5.9.10) in which the velocity of light c is replaced by the average velocity of sound $\langle u \rangle$ (the averaging is carried out by taking into consideration the type of lattice symmetry and the nature of acoustic waves, i.e. the transverse and longitudinal nature of sound). Consequently, the lattice heat capacity at low temperatures (when formulas (5.10.13) and (5.10.10) are valid) can be written in the form

$$C_V = \left(\frac{\mathrm{d}E}{\mathrm{d}T}\right)_V \approx \frac{Vk^3}{(\hbar \langle u \rangle)^3} T^3. \tag{5.10.14}$$

In the reverse limiting case of high temperatures, whose region is defined by the inequality

$$T \gg \frac{\hbar u}{ka}, \tag{5.10.15}$$

the concept of phonons is no longer applicable. As was mentioned above, the classical approximation is applicable in the present case. In this approximation, an energy kT cor-

[7] To a certain extent, we are carrying out only a qualitative analysis here, and hence the analogy between liquids and crystals is acceptable. It should, however, be noted that only normal longitudinal sound, i.e. elastic compression and extension waves, can exist in liquids. In crystals, as was mentioned above, there also exists transverse sound, viz. shear waves.

responds to each vibrational degree of freedom. Hence, at high temperatures, the mean thermal energy of the lattice nuclei is

$$E = 3NkT, \tag{5.10.16}$$

where N is the total number of nuclei in the crystal. Accordingly, the heat capacity is given by the expression

$$C_V = 3Nk. \tag{5.10.17}$$

For one mole of a substance, the high-temperature heat capacity of the lattice is

$$C_V = 3R. \tag{5.10.18}$$

5.10.6. Electron Gas in Metals. The collectivized outer (valence) electrons in a metal are free to move. The electrons are free in the sense that they can rather easily pass from one ion to another in the lattice. The "gas" of free electrons is of quantum nature. This can be seen from the estimate of their degeneracy temperature

$$T_d \sim \frac{\hbar^2}{m_e k}\left(\frac{N_e}{V}\right)^{2/3}. \tag{5.10.19}$$

Substituting the numerical values, we obtain

$$T_d \sim \frac{(10^{-34})^2}{10^{-30}\times 10^{-23}}(10^{28})^{2/3}\ \text{K} = 10^4\ \text{K}.$$

Here, we have used the fact that there is one free electron per atom in typical metals, i.e. the number density of free electrons in a metal is $(N_e/V) \sim 10^{28}$ electrons per cubic metre. Thus, *at all real temperatures, free electrons in metals obey quantum statistics.*

The quantum description of the free electron states in a metal is carried out by specifying three components of momentum (p_x, p_y, p_z) and the projection s_z of the spin on a certain direction (see Sec. 4.9). A precise value of the electron momentum makes the values of its coordinates quite indeterminate in accordance with the quantum relations (4.3.2), i.e. *the electron may be localized at any point in the crystal.*

It is appropriate to compare at this stage the quantum and classical descriptions of a free particle. In classical mechanics, the state of a particle is defined by exactly specifying

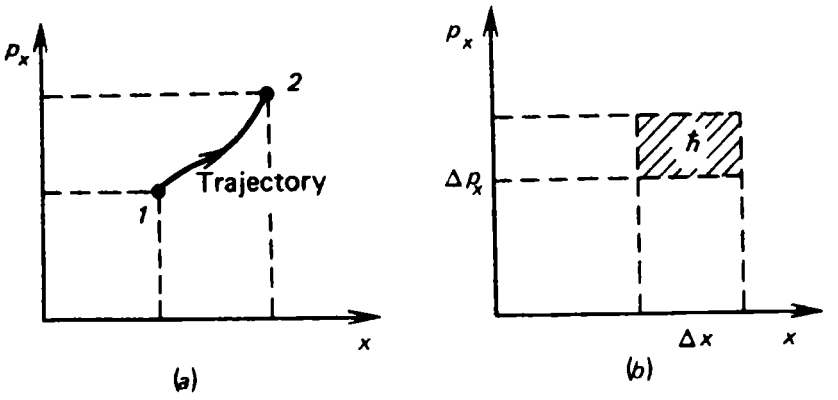


FIG. 5.28

all the three momentum components (p_x, p_y, p_z) and three coordinates (x, y, z). We introduce in a purely formal manner a six-dimensional space of states with axes p_x, p_y, p_z, x, y, z. The state of a classical particle is then defined by a point in this "phase space" (Fig. 5.28*a*). In quantum mechanics, the state of a particle is defined by a cell in the "phase space" in view of the uncertainty relations (4.3.2) (Fig. 5.28*b*). Consequently, a unit cell of states for the quantum description of a free particle in the phase space will be given by

$$\Delta p_x \, \Delta p_y \, \Delta p_z \, \Delta x \, \Delta y \, \Delta z \sim \hbar^3. \qquad (5.10.20)$$

Let us now consider the population rules for unit cells in the phase space. We already know that if particles have an integral spin (including zero), a cell may accommodate an unlimited number of such particles (this applies, among other particles, to photons as well). If particles have a half-integral spin (for example, the spin of the electron is 1/2), not more than two particles may occupy a cell. This is so because the quantum state of electrons with a given momentum (p_x, p_y, p_z) is not defined completely, and the value of one of the two spin projections s_z must also be specified. Thus, in the "momentum space", the cells of quantum states are successively filled by pairs of free electrons.

What is the configuration of the aggregate of all quantum cells in the momentum space? In a "gas" of free electrons (for a cubic lattice), all directions of momentum have nearly the same probability. This means that the "bunch of cells"

will be transformed into a "spherical structure". The radius of the sphere in the momentum space (we denote this radius by p_F) is completely determined by the number density of free electrons N_e/V.

Let us consider a degenerate electron gas at absolute zero. At $T = 0$ K, the energy of any system assumes its lowest possible value under given conditions. Since the energy of free electrons is $\varepsilon = p^2/(2m_e)$, the filling of cells by electron pairs begins from the cell surrounding the origin of momentum, i.e. from the zeroth cell ($p = 0$).

Let us now calculate the number of quantum cells in a very thin spherical layer of thickness dp and with an arbitrary "radius", viz. momentum p (the volume V of a crystal appears as a multiplier in the general expression):

$$\frac{4\pi p^2\, dpV}{\hbar^3}. \tag{5.10.21}$$

The number of electrons filling these cells is double the quantity (5.10.21). Integrating this quantity between $p = 0$ and $p = p_F$, we obtain the total number of free electrons

$$N_e \approx \frac{V}{\hbar^3}\int_0^{p_F} p^2\, dp = \frac{V}{\hbar^3}\left(\frac{p_F^3}{3} - \frac{0}{3}\right) \approx \frac{V}{\hbar^3}\, p_F^3,$$

whence

$$p_F \sim \hbar\left(\frac{N_e}{V}\right)^{1/3}. \tag{5.10.22}$$

The momentum p_F is called the **limiting**, or **Fermi, momentum.**

In order to find the value of the **limiting**, or **Fermi, energy** ε_F, we must know the relation between ε_F and p_F. Recalling that the velocity of an electron in an atom is two orders of magnitude lower than the limiting velocity c (see Sec. 4.5.1), we conclude that the relation between p_F and ε_F must be nonrelativistic, i.e.

$$\varepsilon_F = \frac{p_F^2}{2m_e} \approx \frac{\hbar^2}{m_e}\left(\frac{N_e}{V}\right)^{2/3}. \tag{5.10.23}$$

It should be noted that (5.10.23) is identical to (5.8.10). The limiting energy of free electrons in a metal determines

their degeneracy temperature:

$$kT_d \approx \varepsilon_F. \tag{5.10.24}$$

This means that the formulas for an electron gas in metals obtained in this section are valid not only at $T = 0$ K but also at all temperatures $T \ll T_d$. The degeneracy temperature T_d is of the order of several tens of thousand kelvins (see Eq. (5.10.19)).

At temperatures $kT \ll \varepsilon_F$, only the boundary electrons are excited. The Fermi distribution function is spread over kT about the value ε_F (see Fig. 5.9). Hence the fraction of **Fermi electrons** (i.e. the electrons belonging to the partially filled band) participating in the thermal motion is approximately equal to kT/ε_F. At temperature T, the number of such electrons will be

$$n_e(T) = N_e \frac{kT}{\varepsilon_F}, \tag{5.10.25}$$

where N_e is the total number of electrons in the partially filled band.

Taking (5.10.24) into consideration, we can transform (5.10.25) as follows:

$$n_e(T) = N_e \frac{T}{T_d}. \tag{5.10.26}$$

This is the number of thermal electrons, which increases with temperature. At $T = 0$ K, the number of such electrons is zero. At room temperature, the number of thermal electrons in metals (in a cubic metre) is of the order of magnitude

$$10^{28} \frac{10^2}{10^4} \sim 10^{26}.$$

The number of nuclei in the lattice is about two orders of magnitude larger. Hence, at room temperature, the contribution of electrons in a metal to the values of thermodynamic quantities is just a few percent of the lattice contribution.

Having made these remarks, we can easily write the expression for the electron heat capacity in metals:

$$C_V^e = \frac{dE_e}{dT} = \frac{d}{dT}[n_e(T)\,T]$$

$$= \frac{d}{dT}\left[N_e\left(\frac{T}{T_d}\right)T\right] \approx N_e \frac{T}{T_d}. \tag{5.10.27}$$

For one mole of a substance, we can write

$$C_V^e \approx N_A \left(\frac{T}{T_d}\right) k = \left(\frac{T}{T_d}\right) R. \tag{5.10.28}$$

This is a very convenient form of expression, but it does not "show" that it is applicable to electrons. Using formula (5.10.27) expressing the temperature in energy units, we obtain

$$C_V^e \approx N_e \frac{T}{T_d} = N_e \frac{T}{\varepsilon_F} = \frac{m_e}{\hbar^2} N_e T \left(\frac{V}{N_e}\right)^{2/3}. \tag{5.10.29}$$

This expression has been obtained with the help of (5.10.24) and (5.10.23).

In metals at low temperatures, the lattice component of the heat capacity is $C_V^{lat} \sim T^3$ (see Eq. (5.10.14)), while the electron component $C_V^e \sim T$ (see Eq. (5.10.29)). Hence a decrease in temperature causes a more rapid decrease in C_V^{lat} than in C_V^e. At temperatures close to absolute zero, it is shown by theoretical estimates and by experiment that the electron heat capacity not only becomes comparable with the lattice heat capacity but even exceeds it.

Before concluding this section, let us consider an extreme state of matter, viz. the **ultrarelativistic degenerate electron gas**. For electron energies considerably exceeding their rest energy ($\varepsilon \gg m_e c^2$), the momentum and energy are connected through the relation

$$\varepsilon \approx cp. \tag{5.10.30}$$

This dependence is obtained from (2.7.6) under the assumptions made above. Accordingly, the relation between the limiting energy ε_F and the limiting momentum p_F is given by

$$\varepsilon_F \approx cp_F = \hbar c \left(\frac{N_e}{V}\right)^{1/3}. \tag{5.10.31}$$

In such a state, the electron gas may exist in white dwarfs, i.e. in large dense stars (see Chap. 9 for details).

5.11. Phase Transitions

The equation of state for an ideal gas provides a reliable description of any molecular or atomic gases under pressures that are not far in excess of normal pressures. In compressed gases, when the interaction between molecules becomes noticeable, the thermal properties become quite peculiar, and their behaviour can no longer be described by a single equation of state.

The interaction between particles in a crystal is quite strong, but the nature of their thermal motion is quite simple (small oscillations). Consequently, the thermal dynamic properties of crystalline solids can be analyzed in general form, i.e. in a form that can be used for all crystalline solids. However, such an analysis is possible only at low and high temperatures.

Liquids are substances in which the interaction between particles is strong, but at the same time the thermal motion of the particles is quite complicated. This makes it difficult to theoretically investigate the statistical properties of liquids. The nature of interaction between particles of a liquid, which is different for different liquids, strongly affects the temperature dependence of thermodynamic quantities. Hence the thermal properties of different liquids are quite different. Thermal properties of liquids can be analyzed in general form only at temperatures close to absolute zero. Helium is the only substance in nature which remains in liquid state at $T \to 0$ K.

After a brief review of the state-of-the-art in thermodynamic investigations of a substance, we go over to **phase transitions.** These include the processes of transitions from the crystalline to the liquid state (melting) and from the liquid to the gaseous state (evaporation). The states of a substance between which the transitions take place are called **phases.**

Both phases in a phase transition are in a state of thermal equilibrium characterized by certain values of temperature and pressure (these values are different for different substances). We have already shown that all parts of a system in the state of thermal equilibrium are at the same temperature (see Sec. 5.5). However, thermal equilibrium also suggests mechanical equilibrium, i.e. the absence of intrin-

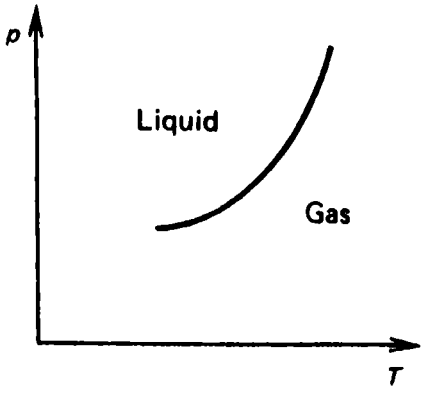


FIG. 5.29

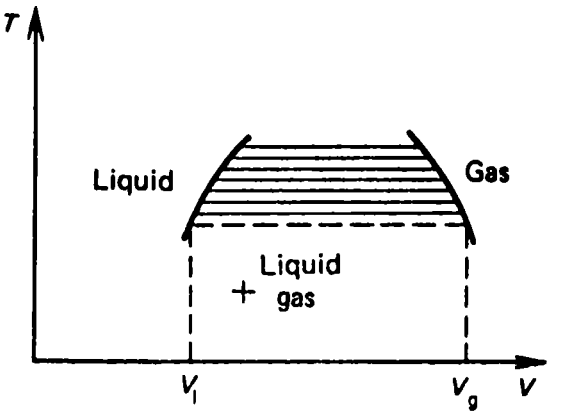


FIG. 5.30

sic macroscopic displacements within the system. This means that the pressure in the system is the same everywhere.

The relation between the pressure p and temperature T of a phase transition is represented in the form of **phase diagrams** (by way of an example, Fig. 5.29 shows the liquid-gas phase transition). If the external pressure changes, the temperature of the phase transition changes accordingly. Different phases are in coexistence at points of the phase diagram. Away from the phase curve, the substance is homogeneous, i.e. in a single-phase state.

Phase transitions can be represented not only in terms of the variables p and T but also T and V or p and V. Indeed, while different phases of a particular substance have their own relations connecting p, V, and T (their equations of state), there is a universal relation connecting these parameters for the phase transition (this relation will be derived below).

Figure 5.30 shows the phase diagram of the liquid-gas transition in terms of variables T and V. A certain amount of substance (1 kg) is usually considered so that V is the specific volume of the substance, and $1/V$ its density. Instead of the phase curve on the p vs. T diagram, we now have an entire phase region for T and V in which the substance may coexist in the gaseous and liquid states in different proportions.

We shall now proceed to derive the thermodynamic relation describing phase transitions. To begin with, we remark that the phase transition is abrupt and is always accompanied by the absorption or liberation of a certain

amount of heat q, called the **specific latent heat of transition.** For the solid-liquid transition, this heat is called the **specific latent heat of fusion,** while for the liquid-gas transition, it is called the **specific latent heat of vaporization.** For the direct (transition from phase *1* to phase *2*) and reverse (from *2* to *1*) processes in the phase transition

	Solid	
(fusion)	↓↑	(solidification)
	Liquid	
(evaporation)	↓↑	(condensation)
	Gas	

the following relation is obviously satisfied:

$$q_{12} = -q_{21}. \tag{5.11.1}$$

In other words, *if heat is absorbed in a certain phase transition, the same amount of heat will be liberated during the reverse transition.*

Let us consider a transition from phase *1* to phase *2* on the p-V diagram in Fig. 5.31. Since the phase transition is an equilibrium process involving the expenditure of specific latent heat q_{12}, we can naturally use the general thermodynamic relation (5.7.5) for obtaining the required relations. Specifically, we assume that we are dealing with a very "narrow" Carnot cycle (see the end of Sec. 5.8.5) with isotherms *ab* and *cd* (see Fig. 5.31). Isotherm *ab* corresponds to a transition from phase *2* to phase *1*, while isotherm *cd* corresponds to a transition from phase *1* to phase *2*. Each of these transitions takes place at a constant pressure, transition *ab* taking place at pressure p and transition *cd* occurring at pressure $p + \mathrm{d}p$. Segments *bc* and *da* of the "Carnot phase cycle" must generally be adiabats, but this is not important in the problem under consideration since we are interested only in the "phase area" enclosed within the contour *abcda*. Since the height $\mathrm{d}p$ of the contour is extremely small, the form of the "lateral sides" of the

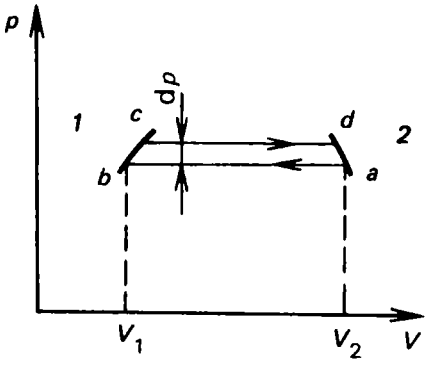


FIG. 5.31

contour will not significantly affect the phase area. The phase area is $(V_2 - V_1)\,dp$, where V_2 and V_1 are the specific volumes of the corresponding phases between which the transition takes place. This area is nothing but the work described by Eq. (5.7.3). On the other hand, this work is equal to the product of the specific latent heat q_{12} of transition (along isotherm *cd*) and the efficiency of the Carnot cycle (see Eq. (5.7.6)). The latter quantity is defined in the present case as dT/T, where dT is the temperature difference between the two isotherms. Consequently, we obtain

$$(V_2 - V_1)\,dp = q_{12}\,\frac{dT}{T},$$

or

$$\frac{dp}{dT} = \frac{q_{12}}{T(V_2 - V_1)}. \tag{5.11.2}$$

As an example, let us apply this formula to the process of evaporation. We shall determine the **saturated vapour pressure** p over the surface of a liquid as a function of the temperature T. A **saturated vapour,** i.e. vapour in equilibrium with its liquid, is usually obtained in a closed vessel. If the temperature is quite low so that the saturated vapour can be treated as an ideal gas, the required dependence can be obtained rather easily (the reader will have the opportunity to get acquainted with the process of **integration** of simple differential equations). We write Eq. (5.11.2) for the case under consideration in the form

$$\frac{dp}{dT} = \frac{q}{T(V_g - V_l)}. \tag{5.11.3}$$

If we consider one mole of a substance, q will represent the molar heat of evaporation, and V_g and V_l the molar volumes of gas and liquid respectively. Usually, $V_g \gg V_l$ (in the case of water, for example, $V_g \simeq 1600 V_l$ at 100 °C). Hence we can neglect V_l in Eq. (5.11.3). Using the condition for ideal saturated vapour, we can express V_g in terms of pressure and temperature in accordance with Eq. (5.8.4). This gives $V_g = RT/p$. Substituting this expression into (5.11.3), we obtain

$$\frac{dp}{dT} = \frac{qp}{RT^2}. \tag{5.11.4}$$

It should be recalled that we have to determine p as a function of T. For this purpose, we transform (5.11.4) to

$$\frac{1}{p}\frac{\mathrm{d}p}{\mathrm{d}T} \equiv \frac{\mathrm{d}\ln p}{\mathrm{d}T} = \frac{q}{RT^2}. \tag{5.11.5}$$

Here, we have used the expression for the derivative of a logarithmic function. At this stage, the physics of the process allows us to make a mathematical simplification. Indeed, the heat of transition has a weak dependence on temperature (for example, the heat of evaporation of water decreases by just 10% as the temperature changes from 0 °C to 100 °C). Hence we can treat q as a constant. As a result, Eq. (5.11.5) can be written in the form

$$\frac{\mathrm{d}\ln p}{\mathrm{d}T} = -\frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{q}{RT}\right). \tag{5.11.6}$$

Here, we have used the relation $\mathrm{d}(1/x)/\mathrm{d}x = -1/x^2$ and considered the fact that a constant can be taken out of the differentiation symbol. Finally, we take the indefinite integrals of both sides with respect to T:

$$\int \frac{\mathrm{d}\ln p}{\mathrm{d}T}\,\mathrm{d}T = -\int \frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{q}{RT}\right)\mathrm{d}T + \text{const}.$$

As a result of integration, we obtain

$$\ln p = \text{const} - \frac{q}{RT}.$$

Finally, we arrive at the expression

$$p = \text{const} \cdot \mathrm{e}^{-q/(RT)}. \tag{5.11.7}$$

This relation shows that *the saturated vapour pressure increases very rapidly (exponentially) with temperature.* It is appropriate to mention at this stage that the curve showing the equilibrium of two phases (gas and liquid) in Fig. 5.29 is nothing but a graphical representation of Eq. (5.11.7).

The following circumstance is also worth noting. The thermodynamic relation (5.11.3) is exact, while (5.11.7) is just an approximate expression. This is obvious from the very approach used for deriving these formulas. Hence if a detailed comparison is made between the empirical phase curve showing the equilibrium of liquid and vapour phases (see Fig. 5.29) and dependence (5.11.7), the two will coincide only in the low-temperature region.

The integration constant in (5.11.7) cannot be determined thermodynamically without resorting to experiment. It would be possible to calculate this constant only if we could theoretically construct a reliable microscopic model for the process of evaporation. This, however, is not an easy task.

Formula (5.11.7) can be obtained directly from simple physical considerations without solving the differential equation (5.11.4). After all, a molecule in a liquid, which is bound to other adjoining molecules, can be transformed to the vapour phase only by doing work. This work is obviously equal to the difference between the potential energies U_g of the molecule in the gaseous phase and U_l in the liquid phase. For one mole of a substance, the amount of work done for a molecule to escape is q/N_A, where q is the molar heat of evaporation, and N_A is Avogadro's number. Using Eq. (5.6.5), we can write

$$\frac{(\text{Density})_g}{(\text{Density})_l} = e^{-(U_g - U_l)/(kT)} = e^{-q/(N_A kT)} = e^{-q/(RT)}. \qquad (5.11.8)$$

Since density differs from pressure in only a constant multiplier kT (see Eq. (5.8.2)), we get the same expression for the ratio of pressures:

$$\frac{p_g}{p_l} = e^{-q/(RT)}. \qquad (5.11.9)$$

It can be seen from this expression that as the temperature increases, so does the saturated vapour pressure, and hence the density of saturated vapour as well. According to Eq. (5.11.8), at a certain temperature T_{cr}, the saturated vapour and liquid will have the same density, and the two will become indistinguishable. This temperature is called the **critical temperature.** Hence at $T = T_{cr}$, the phase curve representing the equilibrium of the liquid and gaseous phases must terminate (Fig. 5.32).

Figure 5.32 shows that while the thermal equilibrium of any two phases (solid-liquid, liquid-gas, solid-gas) of a given substance can be attained at different pairs of values of p and T, the equilibrium of three phases is possible only at one point, called the **triple point** of a substance. It follows directly from here that *the thermal equilibrium of more than three phases of a substance is not possible.*

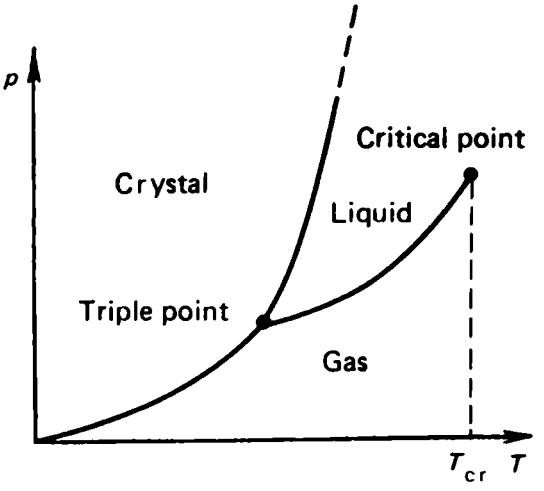


FIG. 5.32

The phase transitions considered above are called **first-order phase transitions.** These transitions are characterized by the fact that $q \neq 0$ for them. However, there are also **second-order phase transitions** for which $q = 0$. As a result of such phase transitions, the state of a substance does not change. In other words, the internal energy E of the substance, its volume V, enthalpy H (see Eq. (5.8.17)), etc. remain unchanged at the moment of transition. However, such a transition is accompanied by an abrupt change in the temperature dependence of these quantities, i.e. the derivatives of these quantities with respect to temperature (dE/dT, dV/dT, dH/dT, etc.), undergo a jump. Thus, the second-order phase transitions affect, among other things, the heat capacities and the thermal expansion of substances. In experiments, the value of the heat capacity C_p is usually measured (see Eq. (5.8.15)). Second-order phase transitions are accompanied by a change in the internal symmetry of bodies.

Examples of second-order phase transitions are:

(i) transition of a metal to the superconducting state;

(ii) ferromagnet-paramagnet transition;

(iii) transition of liquid helium to the superfluid state.

Chapter 6

MACROSCOPIC MOTION OF MEDIA. NONEQUILIBRIUM PROCESSES

6.1. Nonequilibrium States of Bodies

In the previous chapter, we considered the equilibrium states of macroscopic systems and the equilibrium processes taking place in them. By equilibrium processes we meant quite slow processes; it was assumed that the entire system was in a certain equilibrium state at each instant of time, and the process involved a continuous change from one equilibrium state to another.

Quite a slow process is realized when the characteristic time of the process is much longer than the **relaxation time**. The relaxation time determines the time in which the equilibrium is established in the system; if a closed system is in a nonequilibrium state at a certain instant of time, it will most probably start attaining equilibrium at the next moment.

It is important to note that the relaxation time considerably depends on the size of the system and increases with the size of macroscopic bodies. This means that *small parts of a macroscopic system attain equilibrium much sooner than the entire macroscopic body as a whole.* This property of relaxation time allows us to introduce the concept of **local equilibrium**, i.e. equilibrium at a point. A point here means a volume element much smaller than the size of the body, yet containing a large number of molecules or atoms.

In the state of local equilibrium, a "point" in a medium is characterized by its local temperature, while the nonequilibrium medium is characterized by a temperature field. With the passage of time, this partial equilibrium of the closed system is transformed into the total equilibrium, and the temperature of all its parts gradually equalizes.

6.2. Macroscopic Motion

In view of the establishment of local equilibrium at each point in a medium, the equation of state for a given material of the medium is satisfied at each point. Hence, if we know the temperature of different points of the medium (whose density is known), we can simultaneously determine the pressure at these points. In equilibrium systems, the pressure p and temperature T are constant throughout the volume of a body. The constancy of p indicates the mechanical equilibrium of the body, while the constancy of T reflects its thermal equilibrium. Naturally, in the equilibrium state there should be no intrinsic macroscopic motion in the system. In contrast to the microscopic motion of molecules in the system, its macroscopic motion means one in which the macroscopic parts of a body move as a whole. If there are pressure and temperature distributions p (**r**) and T (**r**) in the system, where **r** is the coordinate of an arbitrary point in the medium, the system is not in an equilibrium state. The pressure drop (gradient) in such a body creates intrinsic macroscopic motion of the body characterized by a certain velocity distribution **v** (**r**) of the medium.

In practice, we frequently have to deal with macroscopic motions in liquids and gases. The general laws of such a motion are the same for liquids and gases. In order to describe the macroscopic motion of media in which a whole continuum of points (points adjoining one another) moves instead of limited aggregates of points of the media, we turn (once again!) to the conservation laws. However, it is convenient to apply the conservation laws in **local form** when dealing with continuous media, or continua.

The local formulation of conservation laws is not only convenient and visual but also necessary. Indeed, on account of relativistic principles (see Sec. 2.7) whose main idea is that the velocity of propagation of particles and fields is finite, the following two formulations of the conservation laws are not equivalent (we shall consider the example of the electric charge conservation law):

(i) the total charge in a closed system of moving charged particles is conserved if by this charge we mean the sum of charges "here" and "there" actually counted by us;

(ii) the decrease in charge "here" in a unit time is equal to the flow of charge ensuring its appearance "there".

The second definition is more profound and conveys certain details of the conservation process through the statement that charge is conserved locally.

While considering the macroscopic motion of continua **(hydrodynamics)**, we can use the energy, momentum (angular momentum in some cases), and mass conservation laws. The latter can be interpreted in terms of hydrodynamic concepts as the conservation of the amount of substance (number of particles). These laws of conservation of two scalars and one vector give five equations in all. Accordingly, the state of a moving liquid (or gas) is described with the help of one vector function $\mathbf{v}(\mathbf{r}, t)$ having three components and two scalar functions, say, $p(\mathbf{r}, t)$ and $\rho(\mathbf{r}, t)$, where ρ is the density of the medium. The kinematics of the motion of a liquid is defined by the distribution of its velocity $\mathbf{v}(x, y, z, t)$ which is a function of coordinates x, y, z and time t. The thermodynamic state of the liquid is characterized by the distribution of its pressure $p(x, y, z, t)$ and density $\rho(x, y, z, t)$. Knowing the values of p and ρ, we can determine all the remaining thermodynamic quantities by using the equation of state for the substance.

Before concluding this section, let us consider the limits of applicability of the macroscopic description. Like thermodynamics, hydrodynamics also does not take into consideration the microscopic state of a substance. The media are assumed to be continuous, i.e. their atomic structure is neglected. Hence hydrodynamic equations, like thermodynamic equations, are applicable to any type of medium. The limits of applicability of the hydrodynamic description are quite clearly manifested in the analysis of macroscopic motion of gases. This analysis shows that *the laws of hydrodynamics* (in the present case, gas dynamics) *are applicable as long as the mean free path l of gas molecules* (see below) *is much smaller than the characteristic size L of the system*. For example, in wind tunnels, laws of hydrodynamics are used to study the flow of gases past different types of models at normal pressures. This is so because in this case $l \ll L$, where L is the size of the model. If the pressure in the tunnel is reduced, the flow of the highly rarefied gases will no longer be described by the laws of hydrodynam-

ics since in this case l may turn out to be of the same order of magnitude as L. In general, the assumption $l \to 0$ is used in hydrodynamics.

The limits of applicability of hydrodynamics and thermodynamics are identical. In connection with this remark, it should be recalled that the accuracy of description in thermodynamics is determined by the relative level of thermal fluctuations in the macroscopic system (see Sec. 5.2). The flow of highly rarefied gases must be studied by using the methods of statistical theory of transport phenomena.

6.3. Equations of Hydrodynamics of an Ideal Liquid

A field (for example, an electromagnetic field) and a continuum are identical in the sense that both can be described by specifying the values characterizing them at each point in space or in a medium. Hence we can use the mathematical apparatus of vector fields, introduced earlier in electrodynamics (see Sec. 3.3.1), for the hydrodynamic phenomena as well. After all, we can characterize a liquid through a vector field of velocities where each point in the liquid has a certain velocity vector associated with it. We can introduce the concept of **flux of the vector velocity field** in the same way as in Eq. (3.3.1), replacing the arbitrary vector **A** in the latter equation by the velocity vector **v**. The concept of **circulation of the vector velocity field** is introduced as in Eq. (3.3.2) in which **A** is replaced by **v**. In hydrodynamics, the velocity flux and circulation acquire a clear physical meaning as they can be reduced to the flux and circulation (vortex) of the liquid itself.

6.3.1. Matter Conservation Law in Hydrodynamics. Macroscopic motion in a medium is naturally accompanied by mass transfer. The variation of the quantity $\rho(\mathbf{r}, t)$ obeys the mass conservation law in its local formulation: *the decrease in mass per unit time in a certain volume must be equal to the mass flux through the surface which embraces this volume.* Let us formulate this statement analytically. Let V_0 be the volume under consideration in the bulk of a liquid (we isolate this region mentally). The amount (mass) of the liquid in the volume V_0 is $\langle\rho\rangle V_0$, where $\langle\rho\rangle$ is the

mean density of the liquid in the volume V_0 (after all, the density is also a function of coordinates). The decrease in the mass of the liquid in the volume V_0 per unit time is given by

$$-\frac{\Delta}{\Delta t}(\langle\rho\rangle V_0). \qquad (*)$$

The minus sign has been used in this expression to ensure that this quantity is always positive.

The decrease in the amount of liquid per unit time in the volume V_0 can also be calculated in another way. The volume V_0 is bounded by a certain surface S_0. We introduce a vector area element ds whose magnitude is equal to the area of this element, and the vector is directed along the outward normal to the surface. The mass of the liquid flowing across the area element ds per unit time is $\rho \mathbf{v}\, d\mathbf{s}$ since the scalar product

$$\mathbf{v}\, d\mathbf{s} = |\mathbf{v}|\,|d\mathbf{s}|\cos(\widehat{\mathbf{v}\, d\mathbf{s}})$$

is the volume of the liquid flowing per second. The quantity $\mathbf{v}\, d\mathbf{s}$ is positive if the angle between the vectors v and ds is acute. In case this angle is obtuse, the quantity $\mathbf{v}\, d\mathbf{s}$ is negative. In other words, the elementary flux $\rho \mathbf{v}\, d\mathbf{s}$ is assigned the plus sign if the liquid flows out of the volume V_0 and minus sign if the liquid flows into this volume. The total amount of liquid leaving the volume V_0 in one second can be calculated without imposing any restrictions on the surface area S_0, by summing all the elementary fluxes (with their respective signs). This is the procedure that we should have adopted in the present case. However, we use the mean value of the normal component $\langle v_n\rangle$ of the velocity of flux and write the required expression in the form

$$\langle\rho\rangle\,\langle v_n\rangle\, S_{\text{cl}}. \qquad (**)$$

Finally, comparing Eqs. (*) and (**), we obtain

$$-\frac{\Delta}{\Delta t}(\langle\rho\rangle V_0) = \langle\rho\rangle\,\langle v_n\rangle\, S_{\text{cl}}. \qquad (6.3.1)$$

This is the expression for the **matter conservation law in hydrodynamics.**

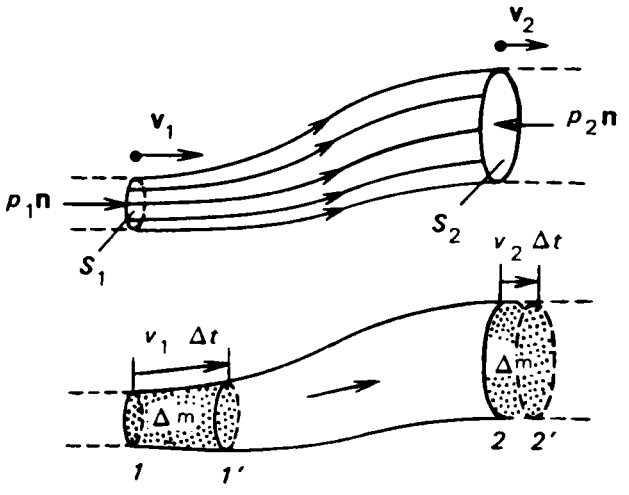


FIG. 6.1

It should be observed that expressions like (6.3.1) are frequently encountered in physics. If a relation of the type (6.3.1) is valid for a quantity and its flux, it means that the quantity is conserved. In the present case, the quantity under consideration is the mass density ρ, while the mass flux density is given by the expression $\rho \mathbf{v}$.

Let us consider the **steady-state flow** of a liquid, i.e. the motion of the liquid in which the quantities describing this motion are independent of time. We use the concept of **a field tube**, viz. a certain volume of the moving liquid enclosed by the **flow lines** (trajectories of the "points of the liquid"). Figure 6.1 shows an example of a field tube. Obviously, the amount of liquid in a field tube remains unchanged, i.e.

$$\frac{\Delta}{\Delta t}(\langle\rho\rangle V_0) = 0.$$

Taking this relation into consideration, we obtain from Eq. (6.3.1)

$$\rho_1 v_1 S_1 = \rho_2 v_2 S_2 = \text{const}, \tag{6.3.2}$$

where S is the cross-sectional area of the field tube at the appropriate places (there is no flow of the liquid across the lateral surface of the tube). This simple but important formula will be put to a number of applications at later stages.

For an **incompressible liquid**, i.e. one whose density along the field tube remains unchanged,

$\rho_1 = \rho_2 = \text{const},$

we obtain the following familiar relation from (6.3.2):

$$\frac{v_1}{v_2} = \frac{S_2}{S_1}.$$

This expression means that the flow of a liquid through a narrow orifice is faster than through a wide one.

6.3.2. Equation of Motion in Hydrodynamics. The momentum conservation law in a liquid must be expressed through a relation of the type (6.3.1), i.e.

$$\frac{\Delta}{\Delta t}\begin{pmatrix}\text{Momentum per}\\ \text{unit volume of}\\ \text{the liquid}\end{pmatrix} = \begin{pmatrix}\text{Momentum}\\ \text{flux}\\ \text{density}\end{pmatrix} S_{\text{cl}}. \qquad (6.3.3)$$

This **equation of motion in hydrodynamics** is analogous to Newton's equation of motion

$$\frac{\Delta}{\Delta t}\mathbf{p}_{\text{m}} = \mathbf{F}, \qquad (*)$$

where $\mathbf{p}_{\text{m}}$ is momentum.

Let us analyze Eq. (6.3.3). It is well known that the force $\mathbf{F}$ has dimensions

$$[F] = [p]\,[S], \qquad (**)$$

where p is pressure. Comparing Eqs. (6.3.3) and (*) and taking (**) into consideration, we conclude that the momentum flux density has the dimensions of pressure. This allows us to write an expression describing the steady-state flow. Indeed, the dimensions of pressure are the same as those of the energy density. Hence it can be expected that the following relation will be satisfied:

$$p + \rho\frac{v^2}{2} = \text{const}. \qquad (6.3.4)$$

Here, we assume that in the case of a stationary liquid ($v = 0$),

$p = \text{const}$

over the entire volume of the liquid. This is true for the case of zero gravity, but we have not taken into consideration the gravitational field.

Keeping in mind the importance of Eq. (6.3.4), we shall give a detailed derivation of this expression. For this purpose, we consider a bundle of flow lines forming the field tube (in Fig. 6.1, n is the unit normal vector). The conservation of mass Δm of the liquid (between cross sections *1* and *2*) during its flow along the field tube means that the amount of the inflowing liquid must be equal to the amount of the outflowing liquid, i.e. $\rho S_1 v_1 \Delta t = \rho S_2 v_2 \Delta t \equiv \Delta m$, where ρ is the density of the liquid, S_1 and S_2 are the cross-sectional areas of the field tube at sections *1* and *2* respectively, v_1 and v_2 are the flow rates in these cross sections, and Δt is the time interval during which the process is investigated. The flow itself is induced by a pressure difference in the liquid. Hence the total work done on the liquid enclosed between cross sections *1* and *2* must be equal to the change in the energy of the mass element Δm of the liquid as it flows from the cross section *1* to section *2*. This follows from the fact that under steady-state conditions, the energy of the liquid enclosed within cross sections *1* and *2* remains unchanged. Consequently,

$$p_1 S_1 v_1 \Delta t - p_2 S_2 v_2 \Delta t = \frac{\Delta m}{2}(v_2^2 - v_1^2),$$

where p is the liquid pressure, and Δm is the mass element under consideration. Since $\Delta m = \rho S_i v_i \Delta t$, where i is the number of the cross section, the last equation is transformed into

$$p_1 + \rho \frac{v_1^2}{2} = p_2 + \rho \frac{v_2^2}{2}.$$

This is just the relation (6.3.4).

The generalization of formula (6.3.4), called the **Bernoulli theorem (equation)**, is expressed through the following equation in a gravitational field:

$$p + \rho \frac{v^2}{2} + \rho\varphi = \text{const}. \qquad (6.3.5)$$

Here, φ is the gravitational potential, and $\rho\varphi$ is the potential energy of the mass element in the gravitational field.

The equation expressing the equilibrium of a liquid in a gravitational field is

$$p(z) = p_0 + \rho g(h - z). \tag{6.3.6}$$

Here, p_0 is the external (say, atmospheric) pressure on the free surface of the liquid at $z = h$. The pressure dependence on only the vertical z-coordinate is natural since a variation of pressure in the horizontal plane would cause a flow of the liquid.

Relations (6.3.2), (6.3.4), and (6.3.5) obtained here contain mechanical parameters only. Hence the motion of a liquid described by these relations must be reversible.

While considering the motion of a liquid in this section, we have not taken into account the internal friction (viscosity) of the liquid and the heat exchange between various moving parts of the liquid. The motion of the liquid in which viscosity and heat conduction are not significant is treated as the **motion of an ideal liquid**. Thus, relations (6.3.2), (6.3.4), and (6.3.5) describe the motion of an ideal liquid.

6.4. Hydrodynamic Analysis of Problems on Viscous Flow, Heat Conduction, and Diffusion

6.4.1. Viscosity. If different "points of a liquid" move with the same velocity in the same direction, the liquid is said to be in equilibrium. Indeed, if we introduce a reference frame attached to the liquid, the latter will be at rest in this reference frame. However, if different parts of the liquid have different velocities, it is not possible to introduce such a reference frame. Such a liquid is in a nonequilibrium state, and spontaneous processes appear in it to bring the system to the state of equilibrium. These processes are responsible for a special property of the liquid, called **viscosity**.

The process of equalization of velocities between adjoining regions of a liquid is accompanied by a momentum transfer. The mechanism of this transfer is of molecular origin and causes an increase in the entropy of the system. Equation (6.3.3) obtained in the previous section describes a purely mechanical *reversible* momentum transfer in the me-

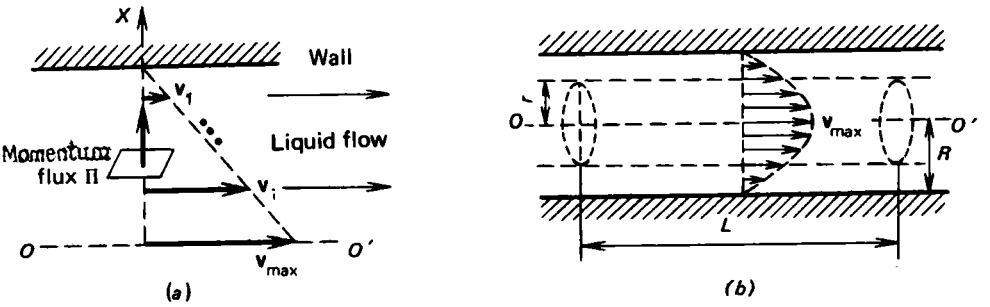


FIG. 6.2

dium. However, if viscosity is taken into consideration, any velocity gradient causes an additional *irreversible* momentum transfer. As the medium moves, the total momentum transfer, including the reversible and irreversible transfer, must satisfy an equation of the type (6.3.3) in accordance with the momentum conservation law.

To write an expression connecting the momentum flux with the velocity gradient, we observe that the flux of a scalar is a vector, while the flux of a vector is a **tensor**. Hence the momentum flux must be a tensor quantity, a tensor being a generalization of the vector concept. Tensor analysis is required for describing those properties of a substance that depend on the direction of a medium.

To avoid the complication of analysis, we shall refrain from using the tensor concept. Mathematics is just a qualitative tool which helps understand the connection between various quantities, "grasp" the main aspects of the phenomenon, and perceive the main conclusions.

Let us consider a plane-parallel flow of a liquid, i.e. its flow in which the velocity vectors $\mathbf{v}_i$ are all in the same direction (Fig. 6.2*a* and *b*). Suppose that the magnitude of velocity v changes (decreases) only along the positive X-direction, i.e. at right angles to the liquid flow. In this case, the velocity is a function of the coordinate, i.e. $v = v(x)$. For small drops (gradients) $\mathrm{d}v/\mathrm{d}x$ of the flow velocity, the momentum flux Π (Fig. 6.2*a*) is proportional to the velocity gradient:

$$\Pi = -\eta \frac{\mathrm{d}v}{\mathrm{d}x}, \qquad (6.4.1)$$

where η is the **dynamic viscosity** of the medium. The momentum flux Π in Eq. (6.4.1) is positive since it is taken along the positive X-direction. The derivative dv/dx, however, is negative since the velocity decreases in this direction. The minus sign in (6.4.1) ensures that the right-hand side of the equality is positive.

Let us find the dimensions of viscosity. Since the momentum flux is the total momentum transferred per unit time across a unit area element perpendicular to the chosen X-direction, we can write

$$[\Pi] = \frac{1\ \text{kg}\cdot 1\ \text{m/s}}{1\ \text{m}^2\cdot 1\ \text{s}} = 1\ \frac{\text{kg}}{\text{m}\cdot\text{s}^2}.$$

The dimensions of the velocity gradient are

$$\left[\frac{dv}{dx}\right] = \frac{1\ \text{m}}{1\ \text{s}\cdot 1\ \text{m}} = 1\ \frac{1}{\text{s}}.$$

Since the dimensions of the left- and right-hand sides of (6.4.1) must be identical, we obtain the dimensions of viscosity:

$$[\eta] = \frac{1\ \text{kg}}{1\ \text{m}\cdot 1\ \text{s}} \equiv 1\ \text{Pa}\cdot\text{s}.$$

All real liquids are viscous. If such a liquid flows, say, along the walls of a cylindrical tube, the thin layer of the liquid adjoining the solid surface "sticks" to the walls on account of the molecular adhesion. Hence the velocity of the liquid flow increases from $\mathbf{v} = 0$ at the walls to a certain maximum value on the axis of the tube. In accordance with the process described by Eq. (6.4.1), a momentum flux is established from the liquid to the walls (the momentum is transferred from faster to slower particles). However, any change in the momentum of the system with time means that, in accordance with Eq. (2.6.1), the system is under the action of a certain force (in the present case, friction). The friction gradually slows down the liquid flow, ultimately transferring its kinetic energy into heat.

6.4.2. Flow of a Viscous Liquid Through a Tube. As an example, let us consider a steady-state flow of a liquid in a cylindrical tube. Let η be the viscosity of the liquid. The tube has a radius R and a length L. The liquid flows along

the axis of the cylindrical tube owing to a pressure drop $\Delta p = p_2 - p_1$ across the tube ends (Fig. 6.2*b*).

The liquid flowing through the tube is subjected to two types of forces, viz. friction which retards the flow and a driving force associated with the pressure drop Δp. In order to write the expressions for these forces, we construct an auxiliary coaxial cylinder of radius r (measured from the cylinder axis). Since the momentum flux Π is directed along the radius from the axis to the cylinder wall, we can write the force of friction in the form

$$F_{fr} = \Pi \begin{pmatrix} \text{Area of} \\ \text{the lateral surface} \\ \text{of the cylinder} \end{pmatrix} = \Pi (2\pi r L).$$

Considering that the momentum flux in this case has the form

$$\Pi = -\eta \frac{dv}{dr},$$

the selected volume of the liquid is subjected to a force of friction from the surroundings:

$$F_{fr} = -2\pi r L \eta \frac{dv}{dr}.$$

The pressure difference Δp acting on the base of the cylinder under consideration gives rise to a driving force

$$F_{pr} = \Delta p \begin{pmatrix} \text{Area of} \\ \text{the cylinder base} \end{pmatrix} = \Delta p (\pi r^2).$$

Since we have a steady-state flow of the liquid, the two forces must be equal ($F_{fr} = F_{pr}$), i.e.

$$-2\pi r L \eta \frac{dv}{dr} = \pi r^2 \Delta p,$$

or

$$\frac{dv}{dr} = -\frac{\Delta p}{2L\eta} r.$$

This is an ordinary differential equation whose solution is obtained as follows. Separating the variables

$$dv = -\frac{\Delta p}{2L\eta} r\, dr$$

and integrating the left- and right-hand sides, we obtain

$$v = -\frac{\Delta p}{2L\eta}\int r\,dr = -\frac{\Delta p}{2L\eta}\left(\frac{r^2}{2}\right) + \text{const.}$$

Here, the arbitrary constant is determined by the boundary conditions: the velocity of the liquid over the surface of the tube, i.e. for $r = R$, vanishes.

The last equation then gives

$$0 = -\frac{\Delta p}{4L\eta}R^2 + \text{const.}$$

Finally, we obtain

$$v = \frac{\Delta p}{4L\eta}(R^2 - r^2).$$

This result means that the velocity in the viscous flow of a liquid varies according to a parabolic law, increasing from zero at the walls to the maximum value $v_{\max} = [\Delta p/(4L\eta)]\,R^2$ on the axis of the cylindrical tube (see Fig. 6.2*b*).

6.4.3. Heat Conduction. A purely mechanical reversible transfer of momentum in a medium naturally entails a reversible transfer of the mechanical energy as well. The expression for the mechanical energy of a unit volume of a liquid appears in the Bernoulli theorem (6.3.5) in which $\rho v^2/2$ is the kinetic energy of a unit volume, p is the work done by the force of pressure on the liquid contained in a unit volume, and $\rho\varphi$ is the potential energy of the mass element under consideration in the gravitational field.

If the temperature of the moving liquid is different at different points, the reversible energy transfer is accompanied by the irreversible transfer of energy from the regions with a higher temperature to the regions with a lower temperature (see Sec. 5.5). The latter energy flow is associated with a transition of the system to the equilibrium state. The process of equalization of temperature, called the **heat conduction**, has a molecular mechanism: molecules of the hotter regions of a medium collide with molecules of adjoining cooler regions during their motion and transfer a part of their energy to these molecules.

It should be noted that the irreversible energy transfer in a flow of a viscous and nonuniformly heated liquid is associated not only with heat conduction but also with friction.

Suppose that the temperature T of a medium changes only along a certain chosen direction which we take as the X-axis. The heat flux is defined as the amount of heat passing through a unit area perpendicular to the X-axis in a unit time. If the temperature drop in the stationary medium is not large, the thermal energy flux density is proportional to the temperature gradient, i.e.

$$j_E = -\lambda \frac{dT}{dx}, \tag{6.4.2}$$

where λ is called the **thermal conductivity**. This quantity determines the rate of heat transfer. By definition, the dimensions of heat flux are 1 J/(m^2·s). Since the dimensions of the temperature gradient are $[dT/dx] = 1$ K/m, thermal conductivity has the dimensions

$$[\lambda] = 1 \frac{\text{J}}{\text{m·s·K}} = 1 \frac{\text{W}}{\text{m·K}}.$$

It can be seen quite clearly from Eq. (6.4.2) that a minus sign between the flux density and the gradient corresponding to it is a must. The derivative dT/dx is negative if the temperature falls in the positive X-direction. Since the heat flow takes place in the direction of decreasing temperatures, i.e. along the positive X-direction, its sign is positive.

6.4.4. Heat Transfer Between Two Walls. Suppose that a layer of thickness d of a substance is bound between two parallel planes having an area S each (Fig. 6.3). The two planes are maintained at temperatures T_1 and T_2 respectively, where $T_1 > T_2$. We shall assume that the thermal conductivity λ of the material is constant. In order to neglect its temperature dependence, we assume that the difference $T_1 - T_2$ is relatively small. Since the heat flow takes place from T_1 to T_2, we take T_1 as the origin and pass the X-axis in the direction from T_1 to T_2.

It is clear from symmetry considerations that the temperature distribution $T(x)$ depends only on x. Let us obtain

an equation for the required temperature distribution. In the expression for the total heat flux

$$Q = j_E S \qquad (*)$$

we use relation (6.4.2). Then Eq. (*) assumes the form

$$Q = -\lambda S \frac{dT}{dx},$$

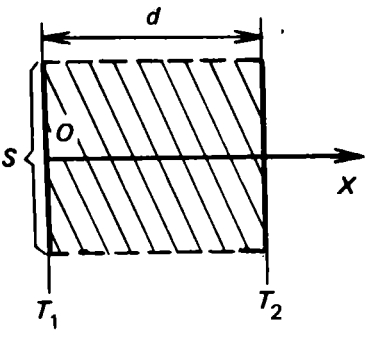


FIG. 6.3

where Q is independent of x.

The differential equation obtained in this way for the required function $T(x)$ must be integrated. The solution will give the temperature distribution in the layer as a function of x. For this purpose, we write the above equation in the form

$$dT = -\frac{Q}{\lambda S} dx.$$

Integrating both sides, we get

$$T = -\frac{Q}{\lambda S} x + \text{const}.$$

It can be seen from here that the temperature is a linear function along the layer thickness.

In order to determine the constant, we must use the conditions at the boundary of the layer. For $x = 0$, the temperature $T = T_1$, and hence the constant is equal to T_1, i.e. the solution is transformed as follows:

$$T = T_1 - \frac{Q}{\lambda S} x. \qquad (**)$$

For $x = d$, the temperature $T = T_2$, and Eq. (**) assumes the form

$$T_2 = T_1 - \frac{Q}{\lambda S} d,$$

or

$$Q = \frac{\lambda S}{d} (T_1 - T_2).$$

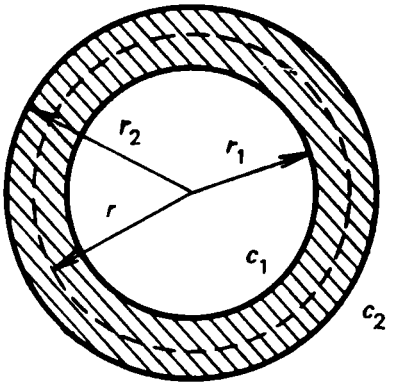


FIG. 6.4

6.4.5. Diffusion. Dissolution of a Solid in a Liquid. So far, we have considered the behaviour of homogeneous liquids. Let us now turn to mixtures of liquids. Suppose that a liquid is made up of a mixture of two substances whose composition is inhomogeneous over its volume. During the macroscopic motion of the mixture, its composition in each part remains unchanged. This conclusion is drawn on the basis of a purely mechanical reversible mass transfer. However, the inhomogeneity of the mixture composition in different parts also leads to an irreversible transfer of the mixture components. This transfer is of molecular origin, and the process of equalization of the mixture concentration is called **diffusion**.

Diffusion is usually studied by taking a uniformly heated stationary inhomogeneous liquid. The composition of the mixture is described by its **concentration** c, defined as the ratio of the mass of one of the constituents of the mixture to the total mass of the liquid in a given volume element. Suppose that a layer of a substance is enclosed between two concentric spherical surfaces of radii r_1 and r_2 (Fig. 6.4). A definite mixture concentration (c_1 or c_2 respectively) of the liquid is maintained on each of these spheres. Suppose that $c_1 > c_2$. The resulting diffusion flux is given by

$$j_D = -D\frac{dc}{dr}, \qquad (6.4.3)$$

where D is the **diffusion coefficient**, or **diffusivity**, dc/dr is the concentration gradient, and the meaning of r is clear from Fig. 6.4. The total flux of the substance is

$$N = 4\pi r^2 j_D,$$

or

$$N = -4\pi D r^2 \frac{dc}{dr}.$$

The last equality can be written in the form

$$\frac{dc}{dr} = -\frac{N}{4\pi D r^2},$$

or

$$dc = -\left(\frac{N}{4\pi D}\right)\frac{dr}{r^2}.$$

Since N is independent of r, the integration of both sides of this equation gives

$$c = \left(\frac{N}{4\pi D}\right)\frac{1}{r} + \text{const}.$$

We now take into account the boundary conditions. For $r = r_1$, the concentration $c = c_1$. Consequently,

$$\text{const} = c_1 - \frac{N}{4\pi D r_1},$$

i.e.

$$c = c_1 + \frac{N}{4\pi D}\left(\frac{1}{r} - \frac{1}{r_1}\right).$$

For $r = r_2$, $c = c_2$. Hence

$$c_2 = c_1 + \frac{N}{4\pi D}\left(\frac{1}{r_2} - \frac{1}{r_1}\right),$$

or

$$N = \frac{(c_1 - c_2)\,4\pi D}{1/r_1 - 1/r_2}. \qquad (*)$$

It follows from Eq. (*) that if $r_2 = \infty$ (an unbounded medium around a sphere of radius r_1), and if the concentration at infinity is zero, we get

$$N = 4\pi D r_1 c_1. \qquad (**)$$

Equation (**) can be used to describe, say, the rate of dissolution of a solid in a liquid, which is determined by the rate of diffusion of the solute in the solvent. Indeed, a saturated solution is immediately formed in the narrow layer surrounding the solid surface. Subsequent dissolution occurs as the solute from this layer diffuses to the surrounding liquid. If the soluble solid has the form of a sphere (of radius r_1), and if the concentration of the saturated solution near sphere is c_1, while the concentration at a large distance

from the sphere is taken equal to zero, the total diffusion flux N from the sphere to the solvent, i.e. the amount of solid dissolving per unit time, is given by Eq. (**).

At this stage, let us determine the dimensions of the diffusion coefficient. If j_D is defined as the number of solute molecules passing through a unit surface area per unit time, we get $[j_D] = 1\ \mathrm{s^{-1} \cdot m^{-2}}$. The concentration is the number of solute molecules per unit volume and has the dimensions $[c] = 1\ \mathrm{m^{-3}}$. Comparing the dimensions of both sides in (6.4.3), we get

$$[D] = 1\ \mathrm{m^2 \cdot s^{-1}}.$$

In order to estimate the characteristic times for diffusion processes, we proceed from dimensional analysis. First of all, *the concentration equalization time in a mixture cannot depend on the concentrations themselves.* Indeed, if the concentration of the mixture changes at different points the same number of times, the diffusion flux will also change accordingly, and the equalization time will remain unchanged. It is obvious that the time τ required for attaining a homogeneous composition of the mixture over the entire system depends only on the diffusion coefficient D and the dimensions L of the region in which different concentrations exist. From these quantities, we obtain the only combination having the dimensions of time:

$$\tau \sim \frac{L^2}{D}. \qquad (6.4.4)$$

This quantity can be called the relaxation time for diffusion processes, i.e. the time in which a homogeneous composition of the mixture is attained over the entire system. It is proportional to the square of the system size and inversely proportional to the diffusion coefficient.

6.5. Kinetic Coefficients in Gases and Their Connection with the Molecular Parameters

Irreversible processes leading to the establishment of equilibrium in a system, such as diffusion, viscosity, and heat conduction, are called **kinetic processes,** and the coefficients associated with these processes are called **kinetic coeffi-**

cients. In liquids and gases, the kinetic coefficients are characteristics of a substance and generally depend on temperature and pressure. The kinetic coefficients are usually determined experimentally, by creating gradients and measuring the corresponding fluxes. By carrying out such measurements at different temperatures and pressures, we can find the dependence of the kinetic coefficients on the state of the substance constituting a certain medium. In liquids and dense gases, there are no general dependences of the kinetic coefficients on temperature and pressure. This is due to the difference in the molecular interaction in different liquids, which is significant for the molecular processes of mass, energy, and momentum transfer.

Ideal gases, i.e. gases whose molecules do not interact (to be more precise, weakly interact) with one another, the relation between the kinetic coefficients and molecular parameters can be theoretically obtained in general form. However, irrespective of how weak the interaction between molecules of such a relatively rarefied gas is, the statistical equilibrium is established in the system only owing to this interaction.

6.5.1. The Concept of Mean Free Path of Molecules. The interaction between molecules (or atoms) separated by rather large distances was analyzed in Sec. 4.10. It was shown that "long-range interactions" between molecules are weakly attractive in nature (van der Waals forces). The interaction between molecules coming quite close to one another is manifested in the form of mutual "impenetrability". The emergence of sharply increasing forces of mutual repulsion between molecules separated by distances of the same order of magnitude as their size is due to two reasons, viz. (i) the Coulomb repulsion between nuclei, and (ii) "quantum-mechanical incompatibility" of outer electron shells of different molecules coming in contact (electrons are Fermi particles, see Sec. 4.6.2).

The dependence of the intermolecular interaction energy U_{12} on the distance r between the centres of molecules is schematically shown in Fig. 6.5. The arbitrariness of this figure stems from the fact that unlike the atomic interaction, molecular interaction depends not only on mutual separation but also on mutual orientation of molecules. In other

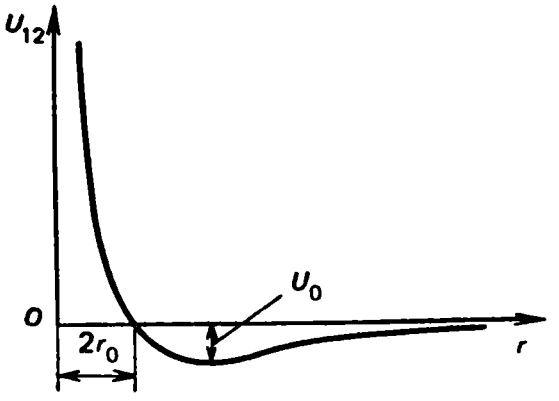


FIG. 6.5

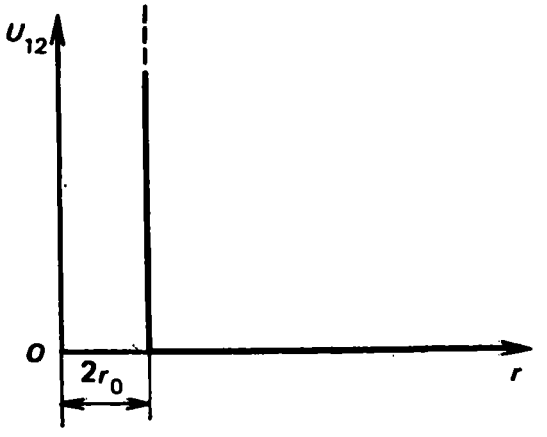


FIG. 6.6

words, U_{12} is a function of many parameters. In this figure, r_0 is the molecular radius, U_0 is the depth of the potential well, which is of the order of magnitude of the critical temperature (see Sec. 5.11) of a given substance, and is a few tenths or hundredths of an electron volt. It should be noted that the analogous function $U_{12}(r)$ for interatomic interaction in a molecule has a minimum U_0 of the order of several electron volts.

Let us consider a real interaction of molecules. This analysis must be carried out quantum mechanically and is quite complicated. In order to simplify this analysis, we go over from the interaction of two molecules (see Fig. 6.5) to the interaction of two atoms (Fig. 6.6). Now r_0 is the radius of an atom, and the interaction is reduced to the brief instants of collisions of the atoms.

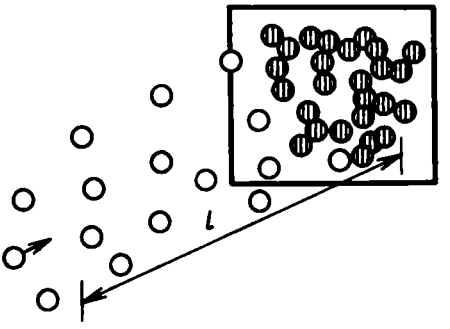


FIG. 6.7

Let us consider a highly rarefied gas of molecules (spheres) in which simultaneous collisions of more than two molecules in a single act of collision can be disregarded. Let n be the number of molecules per unit volume. To find the mean distance l that must be traversed by a molecule to have an appreciable probability of colliding with another molecule, we observe that a thin layer of gas is "transparent" for the molecule under consideration. We project all the molecules in the layer onto a screen of unit area (Fig. 6.7). If the layer has a sufficiently large thickness l, it becomes "opaque" for the molecule under consideration. In this case, nl molecules are projected onto the screen. The fraction of the screen area covered by the projections of moleles is $\pi r_0^2 nl$. If this quantity is close to unity, the layer is opaque. Thus, the distance over which a molecule is almost sure to collide with another molecule is

$$l \sim \frac{1}{\pi r_0^2 n}.$$

The quantity l is called the **mean free path** of a molecule, while

$$\sigma = 4\pi r_0^2 \qquad (*)$$

is called the **effective collision cross section** of molecules. Expression (*) is obtained by considering the fact that the maximum distance between the centres of two spheres (molecules) at which they are still in contact is $2r_0$. Consequently,

$$l \sim \frac{1}{n\sigma}. \qquad (6.5.1)$$

Apart from the quantity l, we can also consider the **mean time** τ between two successive collisions of a molecule. Obviously, in the order of magnitude,

$$\tau \sim \frac{l}{v}, \qquad (6.5.1*)$$

where v is the mean thermal velocity of molecules.

We shall now obtain the numerical estimates for the quantities considered above. The interaction cross section has a value

$$\sigma = 4\pi(10^{-10})^2 \simeq 10^{-19} \text{ m}^2.$$

Under normal conditions, a cubic metre of a gas (at 0 °C and a pressure of 1 atm) contains $n = 2.7 \times 10^{25}$ molecules (see Sec. 5.8.3). Hence the mean free path of molecules has the value

$$l \sim \frac{1}{2.7 \times 10^{25} \times 10^{-19}} \simeq 3 \times 10^{-7} \text{ m}.$$

The mean thermal velocity of (say, nitrogen) molecules is estimated by using formula (5.8.7):

$$v \sim \sqrt{\frac{3kT}{m}},$$

$$v \simeq \sqrt{\frac{3 \times 10^{-23} \times 3 \times 10^2}{30 \times 10^{-27}}} \frac{\text{m}}{\text{s}} \simeq 6 \times 10^2 \text{ m/s}.$$

Hence the mean time between collisions is

$$\tau \sim \frac{3 \times 10^{-7}}{6 \times 10^2} \text{ s} \simeq 5 \times 10^{-10} \text{ s}.$$

The above numerical estimates are in complete accord with the direct experimental measurements (in the order of magnitude) and also with a more consistent theoretical analysis (than the one carried out here). This means that the model of a gas with spherical molecules quite accurately describes the phenomena considered above.

6.5.2. Molecular Treatment of the Diffusion Process. Suppose that the pressure and temperature of a gas are the same over its entire volume. The gas itself is assumed to be a mixture of two components, the composition of the mixture varying in one direction chosen as the X-axis.

We consider one of the gases in the mixture (gas 1). We denote by n_1 the number of molecules of this gas per unit volume and assume that n_1 is a function of the x-coordinate. We construct on the X-axis an arbitrary unit area element with the x-coordinate perpendicular to the X-axis. The number of molecules crossing this area element per second

is $n_1 v$, where v is the mean thermal velocity of molecules. The diffusion flux is defined as the difference in the number of molecules of a given type crossing the area element from left to right and from right to left. Obviously, the flux across the unit area element can be determined by simply taking a layer of gas having a thickness l (where l is the mean free path of molecules) on both sides of the chosen area element. The diffusion flux is then

$$j_D \sim v n_1|_{x-l} - v n_1|_{x+l}. \qquad (*)$$

Here, $n_1|_{x-l}$ is the number of molecules of type 1 in a unit volume in the vicinity of the point $x - l$. The variation of the function n_1 at two neighbouring points $x - l$ and $x + l$ can be expressed as the product of its derivative and the distance between these two points, i.e.

$$n_1|_{x-l} - n_1|_{x+l} \approx -l \frac{dn_1}{dx}.$$

Here, it is assumed that the concentration of molecules of type 1 increases in the positive X-direction. As a result, Eq. (*) assumes the form

$$j_D \sim -vl \frac{dn_1}{dx}. \qquad (**)$$

Comparing Eq. (**) with, say, Eq. (6.4.3), we get

$$D \sim vl. \qquad (6.5.2)$$

Thus, *the diffusion coefficient in gases is equal in the order of magnitude to the product of the mean thermal velocity of molecules and their mean free path.*

Formula (6.5.2) can also be written in another form. Taking into account Eq. (6.5.1), we can write

$$D \sim \frac{v}{n\sigma},$$

where n is the total number of molecules of both species in a unit volume of the gas. Turning to the equation of state (5.8.3) for an ideal gas, we obtain the number density of molecules

$$\frac{N}{V} = \frac{p}{kT},$$

and hence

$$D \sim \left(\frac{vk}{\sigma}\right)\frac{T}{p}. \qquad (6.5.3)$$

Consequently, *the diffusion coefficient is inversely proportional to pressure (at a given temperature)*. The dependence $D(T)$ can be easily obtained if we recall that $v \sim \sqrt{T}$. This gives $D \sim T^{3/2}$.

It should be mentioned here that the interaction cross section σ for real molecules is a decreasing function of temperature since the interaction time decreases with increasing relative velocities of molecules. However, this dependence is quite weak. For spherical molecule model, σ is constant.

6.5.3. Diffusion as a Random Motion of Particles. While investigating the process of diffusion, we inevitably ask ourselves as to why particles are displaced by large distances from their initial position with the passage of time. If the motion of molecules is random and all directions are equivalent, the particle should appear to be "marking time".

Let us consider the motion of a Brownian particle, i.e. the particle that is macroscopic in size and yet small enough to feel the difference in the number of collisions of liquid molecules on its "lateral faces". The disordered and random nature of the resulting action of molecules on the particle leads to an equally disordered and random motion of the particle as a whole. Figure 6.8*a* shows the motion of a Brownian particle observed under a microscope. Each linear segment (AB, BC, CD, DE, EF, . . ., YZ) describes the displacement of the particle from a certain initial position to the final position in equal intervals of time Δt; t denotes the total time for which the observation is made, while N denotes the total number of displacements.[1]

Figure 6.8*b* shows one link in the large polygon consisting of two successive displacements AB and BC. Segment

[1] Naturally, the motion of a Brownian particle observed in the interval between the recording of events is random and chaotic.

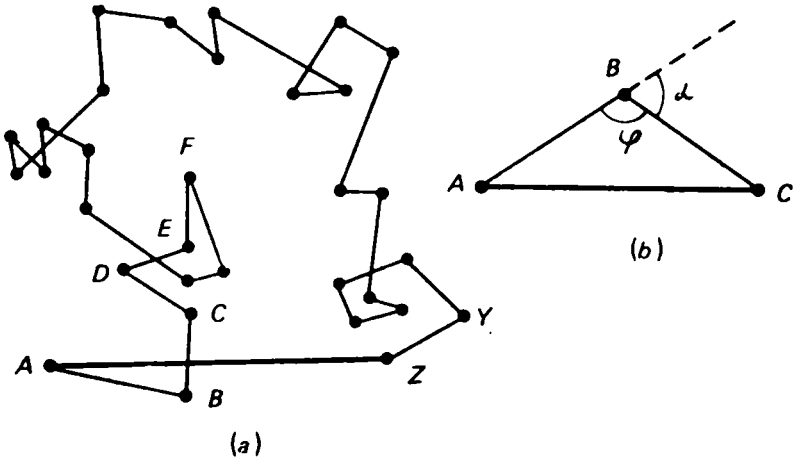


FIG. 6.8

AC is the closing segment. From the well-known geometric theorem, we obtain

$$|AC|^2 = |AB|^2 + |BC|^2 - 2|AB| \cdot |BC| \cos \varphi.$$

Introducing the supplementary angle α to φ and using the equality

$$\cos \varphi = \cos (\pi - \alpha) = -\cos \alpha,$$

we can write the above relation in the form

$$|AC|^2 = |AB|^2 + |BC|^2 + 2|AB| \cdot |BC| \cos \alpha. \qquad (*)$$

Let us now turn to the large polygon. We can assume this polygon to consist of a large number of paired links. For each pair of successive displacements, we can write a relation of the type (*). Adding all these relations termwise and dividing them by the total number N of displacements, we obtain

$$\frac{1}{N} \sum_{i=1}^{N} |A_i C_i|^2 = \frac{1}{N} \sum_{i=1}^{N} |A_i B_i|^2 + \frac{1}{N} \sum_{i=1}^{N} |B_i C_i|^2$$

$$+ \frac{2}{N} \sum_{i=1}^{N} (|A_i B_i| \cdot |B_i C_i| \cos \alpha_i).$$

The last term on the right-hand side containing $\cos \alpha_i$ can be neglected since it is almost equal to zero. This follows from the very nature of Brownian movement which is

random, and hence all the values of the angle α are encountered equally frequently. In other words, terms having almost equal and opposite values have practically the same probability of appearing in this sum.

Observations show that the mean square displacements

$$\langle(\Delta x)^2\rangle = \frac{1}{N}\sum_{i=1}^{N} |A_iB_i|^2 = \frac{1}{N}\sum_{i=1}^{N} |B_iC_i|^2$$

are identical to a fairly high degree of accuracy. This coincidence is the closer, the larger the number N of displacements considered. Consequently, the mean square of the resultant displacement $|AZ|$ (see Fig. 6.8*a*), expressed in the form

$$\langle L^2\rangle = \frac{1}{N}\sum_{i=1}^{N} |A_iC_i|^2,$$

is connected with $\langle(\Delta x)^2\rangle$ through the equality

$$\langle L^2\rangle = N\langle(\Delta x)^2\rangle. \qquad (6.5.4)$$

The number N of displacements is expressed in terms of the total time t of observation and the time interval Δt between successive displacements as follows:

$$N = \frac{t}{\Delta t}.$$

In this case, Eq. (6.5.4) can be presented in the form

$$\langle L^2\rangle = \frac{\langle(\Delta x)^2\rangle}{\Delta t}\, t. \qquad (6.5.5)$$

Note that Eqs. (6.5.5) and (6.4.4) are similar. This similarity is not just quantitative but also appears in the formulation of problems. This leads to the following relation:

$$\langle L^2\rangle \sim Dt. \qquad (6.5.6)$$

This is a remarkable result in many respects. Firstly, it answers the question posed at the beginning of this section as to why a particle moves away from its initial position instead of being confined to a restricted region in space. What is more, relation (6.5.6) describes the "migration" of a particle in time (the "migration" range is proportional

to $\sqrt{t}$). Secondly, the result obtained here reveals the nature of the diffusion process.

6.5.4. Relations Between Kinetic Coefficients. The diffusion flux j_D is associated with the excessive transport of molecules in one direction, i.e. associated with mass transfer. Heat conduction can be treated as a "diffusion of energy", and viscosity as a "diffusion of momentum of macroscopic motion". Since the irreversible transfer of mass, energy, and momentum in a gas involves the same molecular mechanism, the corresponding kinetic coefficients must have the same order of magnitude. We choose the kinetic coefficients that are analogous to one another and play an identical part in the corresponding processes of diffusion, heat conduction, and viscosity. For example, the "diffusion of energy" is described by the **thermal diffusivity**

$$a = \frac{\lambda}{\rho C_p} = \frac{\lambda}{n c_p}, \tag{6.5.7}$$

where λ is the thermal conductivity, C_p is the specific heat at constant pressure, c_p is the heat capacity per molecule, ρ is the mass density, and n is the number density of particles. The "diffusion of momentum of macroscopic motion" is characterized by the **kinematic viscosity**

$$\nu = \frac{\eta}{\rho} = \frac{\eta}{nm}, \tag{6.5.8}$$

where η is the **dynamic viscosity**, and m is the mass of a molecule.

The transport coefficients D, a, and ν have the same dimensions, viz. m^2/s. As a result of the above analysis, we can write

$$D \sim \frac{\lambda}{\rho C_p} \sim \frac{\eta}{\rho}, \tag{6.5.9}$$

or

$$D \sim \frac{\lambda}{n c_p} \sim \frac{\eta}{nm}. \tag{6.5.10}$$

These relations express the connection between kinetic coefficients in gases.

For most real gases, the measured value of the ratio v/a lies between 0.75 and 0.90.

From Eq. (6.5.10), we obtain the following expression for thermal conductivity by taking into account Eq. (6.5.2):

$$\lambda \sim nc_p vl, \tag{6.5.11}$$

while the dynamic viscosity is given by

$$\eta \sim nmvl. \tag{6.5.12}$$

Using expression (6.5.1) for the mean free path l and the relation $v \sim \sqrt{kT/m}$ for the mean thermal velocity of molecules, we obtain from (6.5.11) and (6.5.12) the relations

$$\lambda \sim \frac{c_p}{\sigma}\sqrt{\frac{kT}{m}} \tag{6.5.13}$$

and

$$\eta \sim \frac{\sqrt{mkT}}{\sigma}. \tag{6.5.14}$$

Thus, unlike the diffusion coefficient, thermal conductivity and dynamic viscosity are independent of pressure. The temperature dependence of λ and η can be described approximately as $\propto \sqrt{T}$.

6.6. Resistance to the Motion of Solids in a Liquid

6.6.1. Similitude Method. The Reynolds Number. The general equations of hydrodynamics are complicated on account of their nonlinearity. This is clear even from the simple Bernoulli equation where velocity appears in quadratic form. The situation is aggravated (mathematically) when motion of viscous fluids is studied. Only a few of the viscous-flow problems can be solved exactly. Hence experimental investigations acquire a great significance in hydrodynamics (incidentally, this is true for all branches of physics).

The **similitude methods** are found to be very helpful in hydrodynamic investigations. The essence of these methods is revealed by the following arguments. The only parameter characterizing the liquid itself in the equation of motion for a viscous incompressible liquid is the kinematic

viscosity $\nu = \eta/\rho$. The functions to be determined from a solution of this equation are the velocity $\mathbf{v}$ and the pressure-to-density ratio p/ρ (in the case under consideration, density is constant). The type of a liquid flow is also determined by the boundary conditions and depends on the shape, size L, and velocity u of a body.

We consider a particular type of liquid flow, say, the flow of a liquid past a spherical solid object. We assume the flow to be steady-state, i.e. the flow rate at each point is independent of time. Then the type of motion of a liquid is defined by three parameters, viz. ν, L, and u. Taking into account the dimensions of these parameters

$$[\nu] = 1 \text{ m}^2/\text{s}, \quad [L] = 1 \text{ m}, \quad [u] = 1 \text{ m/s},$$

we form a dimensionless combination. There is only one such combination, called the **Reynolds number**:

$$\text{Re} = \frac{uL}{\nu} = \frac{\rho u L}{\eta}. \tag{6.6.1}$$

The Reynolds number appears in hydrodynamic equations, and its numerical value determines the structure of these equations to a considerable extent.

We introduce the dimensionless coordinates $\mathbf{r}/L$ and the velocity $\mathbf{v}/u$. In other words, we shall express length in units of L and velocity in units of u. Solving the equations of hydrodynamics, we obtain the velocity distribution function in the form

$$\frac{\mathbf{v}}{u} = \mathbf{f}\left(\frac{\mathbf{r}}{L}, \text{ Re}\right). \tag{6.6.2}$$

Let us study the consequences of this relation. If we consider two different flows of the same type (say, the flow of liquids of different viscosities past the spheres of different radii), the "velocity field" (6.6.2) will be the same for both of them only if these flows have the same Reynolds number corresponding to them. The flows can then be obtained from each other by simply changing the scales of measurement of coordinates and velocities. Such flows are called **similar**. This leads to the following statement: *flows of the same type with the same value of the Reynolds number are similar.*

The practical significance of this statement lies in the possibility of simulating real (large-scale) flows in wind

tunnels. Before an aircraft is constructed, its miniature model is subjected to a gas flow whose parameters ensure the equality of the Reynolds numbers for the model and the full-size object.

6.6.2. Drag at Low Velocities. To begin with, let us find how each term in the hydrodynamic equation of motion depends on the Reynolds number. For this purpose, we write down the orders of magnitude of these terms. We proceed from the simplest equation of motion for an ideal liquid. For a steady-state liquid flow, it is the Bernoulli equation

$$\rho u^2 = -p. \qquad (*)$$

Let us now take into account the viscosity of the liquid. Since the irreversible momentum flux is proportional to the velocity gradient, the proportionality factor being the dynamic viscosity, relation (*) can be written in the form

$$\rho u^2 - \eta \frac{u}{L} = -p. \qquad (6.6.3)$$

The notation used here is the same as in the last section (see Sec. 6.6.1). The term containing the density ρ is called the **inertial term**, while the **viscosity term** contains the coefficient η. It can easily be seen that the Reynolds number is given by the ratio

$$\mathrm{Re} = \frac{\text{Inertial term}}{\text{Viscosity term}}. \qquad (6.6.4)$$

For low values of the Reynolds number ($\mathrm{Re} \ll 1$), the inertial term is small in comparison with the viscosity term (see Eq. (6.6.4)). In accordance with Eq. (6.6.3), the velocity is also low. Hence at low velocities of a liquid flow, the inertia of the liquid flowing past a body need not be taken into account. In this case, the flow lines in the inflowing flux (say, onto a cylinder) have the form shown in Fig. 6.9*a*.

Let us derive an expression for the drag force F exerted by a liquid on a body having a characteristic size L and moving with a low velocity u. The properties of the liquid in this case can be described in terms of the dynamic viscosity η alone.

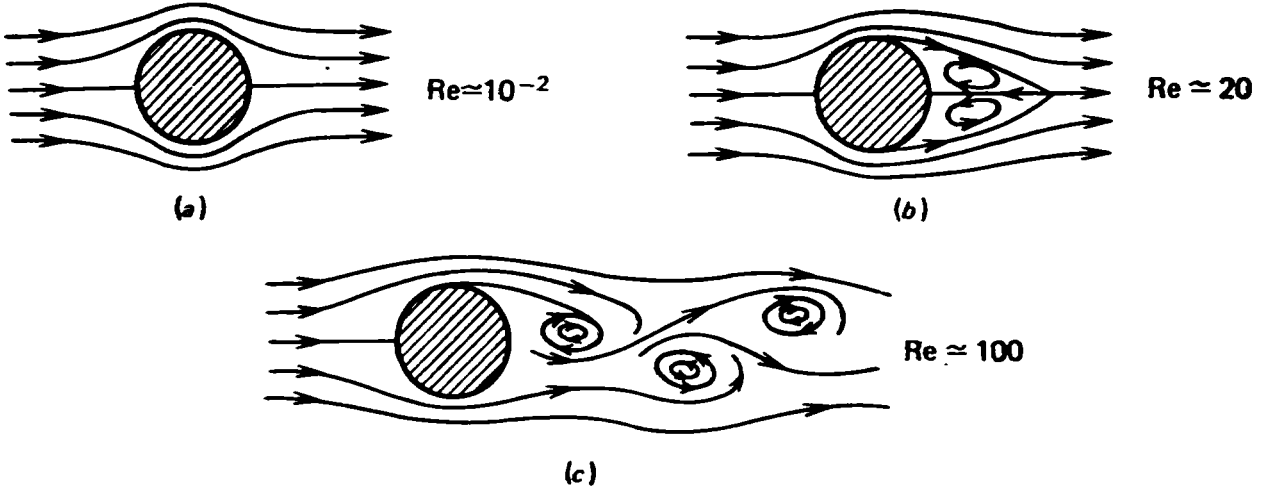


FIG. 6.9

The dimensions of the quantities involved in this problem are:

$$[F] = 1\,(\text{kg}\cdot\text{m})/\text{s}^2,\ [\eta] = 1\ \text{kg}/(\text{m}\cdot\text{s}),\ [L] = 1\ \text{m},\ [u] = 1\ \text{m/s}.$$

It can be seen from here that the product ηLu has the same dimensions as force, i.e.

$$F \sim \eta Lu. \tag{6.6.5}$$

This is the required formula.

For a sphere, the numerical factor in (6.6.5) is 6π. The exact expression in this case has the form

$$F = 6\pi\eta Ru, \tag{6.6.6}$$

where R is the radius of the sphere. This relation is called **Stokes' law**, and is valid for slow motion of a sphere in a liquid.

Stokes' law is found to be helpful in a number of cases. It can be used to determine the velocity of macroscopic particles (which can approximately be treated as spheres) moving in a liquid under the action of a certain force. For example, the settling of particles takes place under the force of gravity, separation of particles by mass in a centrifuge takes place under the action of centrifugal forces, and nonuniformly distributed suspended particles in a liquid participate in diffusion processes under random fluctuations of pressure.

Millikan's classical experiments on the determination of the elementary electric charge were also based on Stokes' law. Let us recollect the technique of Millikan's experiments. Minute oil droplets were sprayed into a chamber containing two horizontal plates of a parallel-plate capacitor. The droplets carried an electric charge (imparted to them during atomization or as a result of absorption of ions from the surrounding air). The velocity u of the uniform sedimentation of a droplet was determined with the help of a microscope (no electric field was applied at this stage). The dynamic viscosity η of air, the density ρ of oil, and the strength g of the gravitational field were assumed to be known. In this case, the condition of equality of forces $F_{\text{grav}} = F_{\text{drag}}$, i.e.

$$mg \equiv \frac{4}{3}(\pi R^3 \rho g) = 6\pi\eta uR,$$

was used to find the radius of the droplet:

$$R = \sqrt{\frac{6\pi\eta u}{(4/3)(\pi\rho g)}}.$$

After this, an electric field E was switched on to neutralize the action of gravity. In other words, conditions were created such that the following relation was satisfied:

$$F_{\text{elec}} = F_{\text{grav}} = F_{\text{drag}}.$$

It follows hence that

$$QE = 6\pi\eta Ru.$$

This relation was used to determine the electric charge Q of a droplet. It was found to be an integral multiple of a certain charge

$$e = \frac{Q_n}{n},$$

where $n = 1, 2, 3, \ldots$. The charge e was taken as the elementary electric charge.

6.6.3. Drag at High (Subsonic) Velocities. Let us again turn to Eq. (6.6.4), which shows that for large values of the Reynolds number ($\text{Re} \gg 1$), the inertial term (containing the density ρ) plays a leading role, and the viscosity term is found to be small in comparison with it. In this case, according to Eq. (6.6.3), the velocity values are quite large.

Let us write an expression for the drag offered by a liquid medium to a body moving in it at a rather high velocity. It was mentioned above that the mass density ρ of the liquid assumes a considerable significance in this case. Thus, we have three parameters at our disposal: the size L of the body, its velocity u, and the density ρ of the liquid. From dimensional analysis, we obtain the following expression for the drag force:

$$F \sim \rho u^2 L^2. \tag{6.6.7}$$

The expression for drag at high velocities is markedly different from the corresponding expression (6.6.5) for drag at low velocities. The change in the dependence is caused by a change in the nature of the liquid flow past a

body. As the velocity increases, so does the flow inertia. As a result, the flow may be separated from the body, and eddies may be formed.

This situation is graphically illustrated in Figs. 6.9*b* and *c*. It can be seen from Fig. 6.9*b* that for Re $\simeq$ 20, a countercurrent flows at the back of the body. This is due to the fact that the streams are slightly broadened at the back of the body, and hence the pressure in them increases in accordance with the Bernoulli theorem. It is also clear from this figure that the emergence of eddies is facilitated by the simultaneous existence of currents and countercurrents in some regions. Upon a further increase in the flow rate at Re $\simeq$ 100 (Fig. 6.9*c*), the nature of flow changes once again. The eddies are separated from the body and move with the flow. New eddies are formed in the wake of the body.

For large values of the Reynolds number, the drag depends considerably on the shape of the body. This can easily be understood with the help of Fig. 6.9*b*. Instead of the cylinder with a circular cross section, let us consider a body whose cross section is circular at one end and tapers down at the tail end. The tail occupies the region where two eddies and a countercurrent coexist. In this case, the flow is smoothly in contact with the tail of the body. The critical situation (separation of the flow from the body) will now arise at much higher velocities. Bodies having such a shape are called **streamlined bodies.**

The quantitative dependence of drag on the shape of bodies is expressed by introducing the **drag coefficient** C into Eq. (6.6.7). This gives

$$F = C\rho u^2 L^2. \tag{6.6.8}$$

Here, L^2 stands for the cross-sectional area of the body. For streamlined bodies, the drag coefficient C has values around 0.03-0.05, while for nonstreamlined bodies C = 1.0-1.5.

Formula (6.6.8) defines the drag of bodies moving at relatively high velocities. However, these velocities must be much lower than the velocity of sound in the medium in which the body moves. In this case, there is no need to take into account the compressibility of the medium, i.e. the change in its density. In formula (6.6.8), it is assumed that ρ = const.

6.7. Instabilities in Hydrodynamics

6.7.1. Transition from Laminar to Turbulent Flows. Let us begin by considering an example of instability. For this purpose, we turn to the experiment. We take a vessel and partially fill it with a light liquid (say, water) and then carefully fill it up with a heavier liquid (say, concentrated sulphuric acid). The system may be in equilibrium in the state when the heavier liquid is on top of the lighter liquid (Fig. 6.10). However, this equilibrium will be unstable from the point of view of energy. After all, the total potential energy of the system will be lower if the heavier liquid is at the bottom and the lighter liquid is at the top (the part of the energy released in this way will be spent on raising the lighter liquid above the heavier liquid). Hence even a slight perturbation is sufficient to initiate such a rearrangement of the system.

This was an example of the unstable equilibrium of liquids. Let us now turn to the study of unstable flows of liquids and gases. Even if we have obtained an analytic solution of the hydrodynamic equations for some problem, it does not mean that this solution can be realized in actual practice. Small perturbations are inevitably present in the flow of a liquid, and hence the obtained solution (even if it is an exact one) can be realized only if it is stable to the small perturbations.

Experiments show that as the Reynolds number increases, it attains a critical value Re_{cr} for which a steady-state flow becomes impossible because of its instability with respect to the minutest perturbations. Naturally, each type of motion has its own value of Re_{cr}.

Let us study the nature of the transient regime resulting from the absolute instability of the steady-state flow for $\mathrm{Re} \geqslant \mathrm{Re}_{cr}$.

As a rule, at first periodic flow pulsations are superimposed on the steady-state motion in this case. As the value of the Reynolds number increases fur-

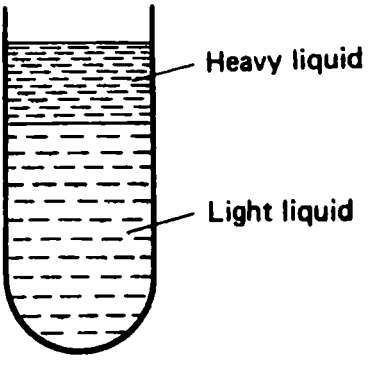


FIG. 6.10

ther, the motion becomes more and more complex and muddled. Such a motion is called **turbulent flow** to distinguish it from the ordered layer-by-layer motion of a liquid, called **laminar flow.**

The turbulent flow of liquids with large values of the Reynolds number is characterized by the existence of an extremely irregular and disordered variation of velocity with time at each point of the flow **(developed turbulence).** The velocity continuously fluctuates about a mean value, the amplitude of this fluctuation may not be small in comparison with the value of the velocity itself. Such an irregular variation of velocity is observed at all points of the flow at a given instant of time. The trajectories of "liquid particles" in the turbulent flow are very complicated, and this causes a strong mixing up of the liquid.

In order to demonstrate the difference between laminar and turbulent flows, we carry out the following experiment. Suppose that water is flowing through a glass tube. Brightly coloured suspended particles are introduced into the flow at the same place of the cross-sectional area of the tube. In the case of a laminar flow, all suspended particles move along the same trajectory. In a turbulent flow, particles are found to have different trajectories. Moreover, it is impossible to predict these trajectories. The trajectories of particles moving in a turbulent flow are random, and particles may appear at any point in the tube.

The above analysis shows that the theory of turbulent flow must be of statistical origin.

6.7.2. Boundary Layer. For large values of the Reynolds number, the viscosity of freely flowing liquids need not be taken into consideration. In the layer near the wall or near the surface of a solid, however, the flow formation depends on the forces of viscous friction. Indeed, the velocity of "liquid particles" at the surface of a stationary solid must be zero. The entire increase in the velocity from zero to its value in the main flow takes place in a thin layer surrounding the body and called the **boundary layer.**

An increase in the velocity of the main flow causes an increase in the velocity gradient in the boundary layer. The laminar flow is accompanied by eddies. As a matter of fact, it is a sharp drop in velocities that is responsible

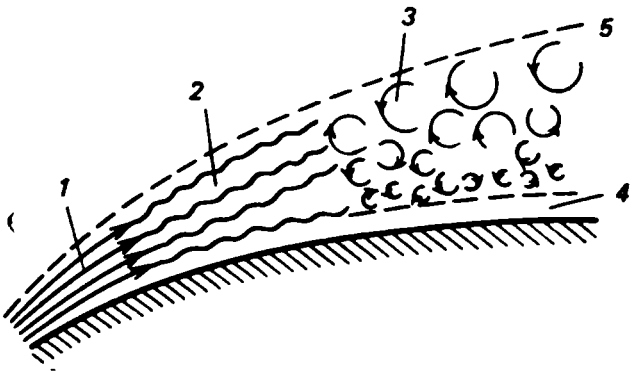


FIG. 6.11

for the emergence of eddies (turbulization) in the flow. These eddies then diffuse to the adjoining region. Figure 6.11 shows the flow structure in the boundary layer (here, *1* is the laminar region, *2* the transition region ("periodic pulsations"), *3* the turbulent region, *4* the viscous sublayer, and *5* an arbitrary boundary between the ideal liquid and the boundary layer).

It should be observed that for $\mathrm{Re} < 420$, the laminar flow in the boundary layer is completely stable to external perturbations. The transition from laminar to turbulent flows occurs at $\mathrm{Re} > 420$. The larger the Reynolds number, the thinner the boundary layer. This is quite clear physically since in the case of an ideal liquid (when the viscosity $\eta = 0$), the boundary layer should be missing altogether.

Thus, an increase in the velocity of the main flow causes an increase in the velocity gradient in the boundary layer, and the molecular momentum transfer ultimately becomes unstable. Momentum transfer now takes place through another mechanism, viz. by macroscopic turbulization of the layer.

The turbulent mixing of a liquid is a much more effective mechanism of momentum transfer than the "diffusion" process of its molecular transfer. This can clearly be seen from the velocity profiles of the laminar and turbulent flows in a tube (Fig. 6.12). In the laminar flow, the momentum transfer takes place through the molecular mechanism, and the velocity of flow varies smoothly from the axis to the walls of the tube. In the turbulent flow, an intensive macroscopic momentum transfer takes place between different regions of the liquid, and this ensures a constant veloc-

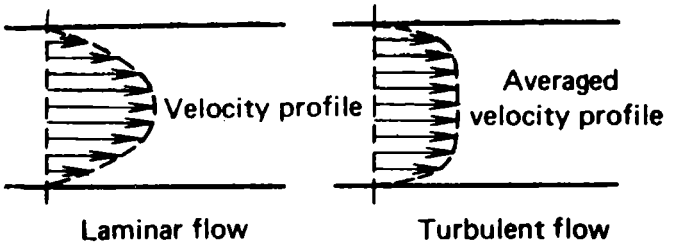


FIG. 6.12

ity of the flow over the entire cross section with the exception of a thin layer near the boundary.

6.7.3. Turbulent Viscosity and Thermal Diffusivity. The statistical nature of the turbulent flow of a liquid makes it possible to draw a certain analogy between this flow and random fluctuations of molecules. The role of the disordered "molecular trajectories" is now played by the disordered "flow lines in the liquid".

It was mentioned above that if a turbulent flow is observed from a certain fixed point, the magnitude and direction of velocity are found to vary randomly.

We shall use the concept of mean velocity $u = \langle v \rangle$ of flow at a point, defined as the time-averaged true velocity v. Then the velocity u must vary smoothly along the flow. The difference

$$v' = v - u$$

between the actual and mean velocities is called the "pulsating" part of the velocity.

Thus an irregular pulsating flow is superimposed on the mean flow.

We know that the kinetic coefficients of molecular transfer of momentum and energy satisfy the relations

$$\nu \sim a \sim \langle v \rangle l,$$

where $\langle v \rangle$ is the mean thermal velocity of molecules, and l is their mean free path. These transport coefficients are used in the laminar flow.

In a flow with developed turbulence, the viscosity and heat transfer may be qualitatively defined by the analogous molec-

ular coefficients. For example, the **turbulent viscosity**

$$\nu_{\text{turb}} \sim L\,\Delta u,$$

where L is the basic scale of turbulence reflecting the size of the region of turbulent flow, and Δu is the change in the mean velocity at distances of the order of L. The relation between the turbulent viscosity ν_{turb} and the molecular viscosity ν has the form

$$\nu_{\text{turb}} \sim \nu \frac{\text{Re}}{\text{Re}_{\text{cr}}}.$$

Accordingly, the heat transfer in the turbulent region will be described by the **turbulent thermal diffusivity**

$$a_{\text{turb}} \sim L\,\Delta u,$$

while the molecular thermal diffusivity a and the turbulent diffusivity a_{turb} are connected through the relation

$$a_{\text{turb}} \sim a \frac{\text{Re}}{\text{Re}_{\text{cr}}}.$$

Thus, molecular and turbulent transport coefficients become comparable in the order of magnitude for critical values Re_{cr} of the Reynolds number.

6.7.4. Transition from Molecular to Convective Heat Transfer. Solar Granulation. Let us now turn to the question of heat transfer in liquids under gravitational fields for very high temperature gradients. At rather low temperature gradients, molecular heat transfer occurs through conduction. However, the molecular mechanism becomes unstable for large values of temperature gradients, and a macroscopic heat transfer by convection takes place.

Suppose that a layer of gas or liquid is enclosed between two horizontal surfaces under gravity. The temperatures of the upper and lower surfaces are T_1 and T_2 respectively; it is assumed that $T_2 > T_1$.

If the temperature difference $T_2 - T_1$ is not too large, the liquid remains at rest, and heat transfer occurs through the usual mechanism of heat conduction. This stage of the process is characterized by the parameters a, L, and $T_2 - T_1$, where L is the thickness of the layer under consideration.

For a certain critical value of the temperature difference $(T_2 - T_1)_{cr}$, the process of molecular heat transfer becomes unstable, and a steady-state macroscopic convective flow sets in. In this case, we must also take into consideration such characteristics of the moving medium as its kinematic viscosity ν. However, no parameter having the dimensions of velocity should appear in the expression for a free convective flow. This is so because the entire motion of the liquid under free convection is caused only by its nonuniform heating. Hence we introduce the **temperature coefficient of volume expansion** β of the liquid, defined as

$$\beta = -\frac{1}{\rho}\frac{d\rho}{dT}.$$

Since the density ρ of the medium is a decreasing function of temperature, we get

$$\rho(T_1) > \rho(T_2).$$

This means that the upper layers are found to be denser than the lower layers. In other words, the system will be unstable in the field of gravity g (see the example at the beginning of Sec. 6.7.1).

Thus, we have the following parameters at our disposal:

$$[a] = [\nu] = 1 \text{ m}^2/\text{s}, [T_2 - T_1] = 1 \text{ K}, [L] = 1 \text{ m},$$

$$[g] = 1 \text{ m/s}^2, [\beta] = 1 \text{ K}^{-1}.$$

From these quantities, we can form two independent dimensionless combinations. The following combinations are usually chosen:

$$\text{Pr} = \frac{\nu}{a}$$

and

$$\text{Gr} = \frac{g\beta L^3 (T_2 - T_1)}{\nu^2}.$$

These parameters are called the **Prandtl number** Pr and the **Grashof number** Gr.

The Grashof number Gr is an important characteristic of the convective flow of a liquid. For small values of Gr, the convective mechanism of heat transfer is absent, and heat is transferred by molecular conduction. For large values of

Gr, the convection becomes turbulent. This means that the role of the Reynolds number Re, which does not appear in the case of free convection (due to the absence of parameters having the dimensions of velocity), is assumed by the Grashof number Gr.

If the boundary conditions for convection are chosen in such a way that the upper surface of the liquid is free and maintained at a constant temperature, the critical value of the temperature difference $(T_2 - T_1)_{cr}$ is attained for

$$\mathrm{Gr} \cdot \mathrm{Pr} = \frac{g\beta L^3 (T_2 - T_1)}{\nu a} > 1100.$$

The convective flow arising in this case is divided into separate regions, or cells, which are reminiscent of the honeycomb structure (Fig. 6.13*a*). The convective flow within each cell of the honeycomb is shown in Fig. 6.13*b*.

For Gr $\sim$ 50 000, the steady-state laminar convection becomes unstable, the honeycomb structure disappears, and turbulent convection sets in.

The cellular structure of a convective flow is observed on a large scale on the granular surface of the Sun. The solar disc contains about 2×10^6 granules (cells) in all. During the period of maximum solar activity, this number slightly increases. The granules have a mean diameter of about 700 km and a mean lifetime of about 8 minutes. The velocity of the ionized gas in the cells attains values of up to 2-3 km/s.

Convective flows are responsible for the transformation of a part of the energy liberated by the inner parts of a star

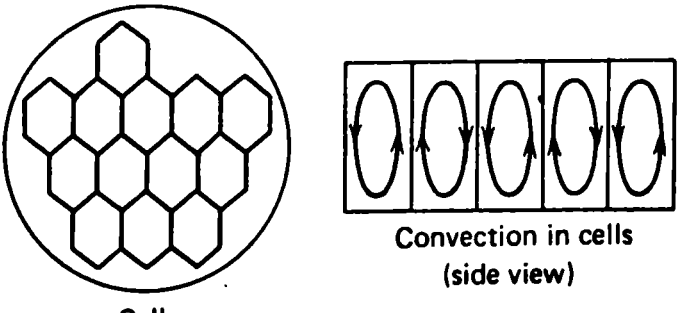


FIG. 6.13

into mechanical or magnetic energy. For example, local magnetic fields may be produced by plasma turbulent regions emerging due to unsteady-state convection (this phenomenon is termed a **turbulent dynamo**).

6.8. Oscillations and Waves in a Liquid

6.8.1. Various Forms of Wave Motion. Any perturbation in quite dense media does not remain localized. On account of an interaction between particles, it propagates in the form of a wave along the entire medium. A wave is a sort of relay transfer of perturbation from one part to another.

Compressive (tensile) and shear strains in crystals play the role of perturbations. These two types of strains can produce elastic (sound) longitudinal or transverse waves in a crystal. Transverse waves are associated with shear strain; in this case, the direction of wave propagation is perpendicular to the direction in which the particles move. Longitudinal waves are associated with compressive (tensile) strain; in this case, the direction of wave propagation coincides with the direction of motion of the particles (parallel or antiparallel).

Elastic shear strain does not appear in a liquid. On the one hand, this fact is responsible for the emergence of flow, and on the other hand, only elastic longitudinal waves (longitudinal sound) appear in the liquid.

However, it should be borne in mind that the properties of the liquid surface are different from the properties in the bulk. The liquid surface is subjected to several forces, viz. the surface tension forces and the force of gravity. The resulting surface (gravitational) waves cause periodic changes in the surface itself. The liquid surface tends to assume an equilibrium plane horizontal shape under the action of gravity. The surface tension forces lie in the tangential plane, while the force of gravity always acts in the vertical direction. Hence the liquid particles in a surface wave perform neiher transverse nor longitudinal vibrations, but rather a circular motion. This motion may become more complicated depending on the conditions under which the waves pass over a given surface. Here, we are speaking of **deep** and **shallow water**.

6.8.2. Wave Characteristics. A sine wave is the simplest form of wave motion (Fig. 6.14). This type of wave is characterized by its **amplitude** A, i.e. the span of the oscillations, the **angular frequency** ω, i.e. the number of cycles over 2π units of time (expressed in radians/s), and the **phase** φ defined as the angular separation of any point on the wave from the zero point (expressed in radians). Sometimes, the height $2A$ of the wave is also used.

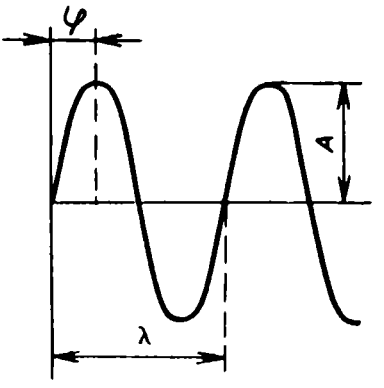


FIG. 6.14

The following statement exists in mathematics: *any function may be presented as a superposition of sinusoids; if this function is periodic, the sinusoids form a discrete set, and their frequencies are multiples of one another; for aperiodic functions, the decomposition into a frequency spectrum is continuous, i.e. the frequencies may assume any values.*

Figure 6.15 shows the superposition of two sinusoids with nearly equal frequencies ω and $\omega + \Delta\omega$, resulting in the formation of a wave packet.

Since a wave is nothing but oscillations propagating in a medium, the **velocity of propagation** is also a parameter characterizing a wave. If we know the frequency ω and the (phase) velocity u_{phase} of propagation of a wave, we can define the **wavelength** as

$$\lambda = \frac{2\pi u_{\text{phase}}}{\omega}.$$

In many media, the velocity of propagation depends on the wavelength; this dependence is called **dispersion**. In such cases, we must distinguish between the **phase velocity** u_{phase} and the **group velocity** u_{group} of wave propagation. The difference between the two becomes especially clear if we consider the propagation of a localized wave packet consisting of a very large number of sine waves with nearly equal frequencies. Such a wave packet spreads (broadens) out on account of dispersion since different sinusoids constituting this packet have different velocities. The velocity of motion of the centre of the wave packet is called the group velocity. This velocity determines the rate of energy trans-

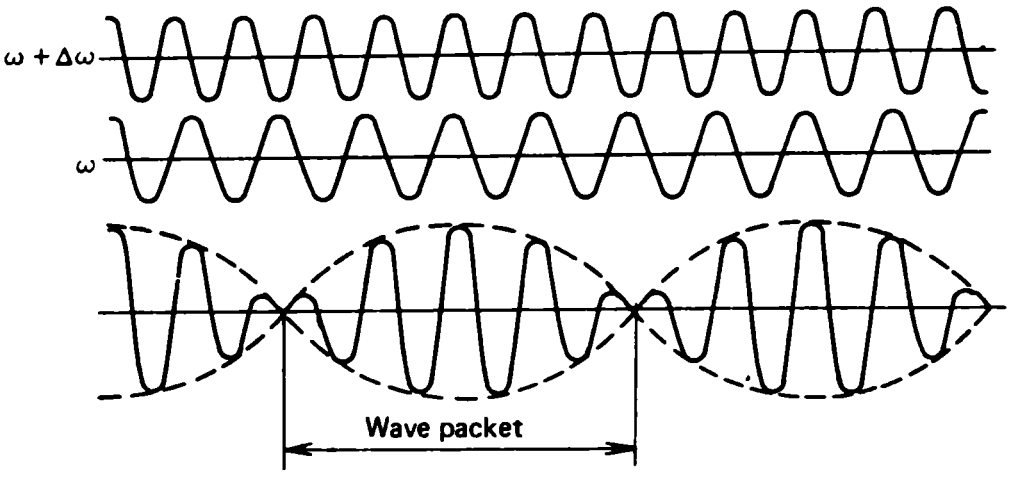


FIG. 6.15

fer by a wave. The velocity of propagation of, say, one of the crests of the wave packet is called the phase velocity. It is the dependence of the phase velocity on the wavelength that determines the dispersion of a wave packet.

Another example showing the difference between the phase and group velocities of wave propagation is shown in Fig. 6.15. Here, we have two sinusoids with nearly equal frequencies ω_1 and ω_2, and **wave numbers** k_1 and k_2. The wave number k is defined as the magnitude of the wave vector **k** coinciding with the direction of wave propagation so that

$$k = \frac{2\pi}{\lambda}.$$

In other words, the wave number shows the variation of phase with distance (radians/m). The phase velocities of the two sine waves are given by

$$u^1_{\text{phase}} = \frac{\omega_1}{k_1}, \quad u^2_{\text{phase}} = \frac{\omega_2}{k_2}.$$

The group velocity is the velocity of the modulated wave obtained by summing the initial waves:

$$u_{\text{group}} = \frac{\omega_1 - \omega_2}{k_1 - k_2},$$

or, in the general case,

$$u_{\text{group}} = \frac{d\omega}{dk}.$$

It is obvious that an infinite monochromatic sine wave is not capable of transmitting a signal on its own. For this purpose, we require a local distortion (modulation) of this wave. The velocity of transmission of this distortion is determined by the group velocity.

6.8.3. Linear and Nonlinear Waves. If the sine waves obtained during the decomposition of a complex wave are independent, i.e. do not influence one another in the course of their propagation, the wave processes are called **linear**. However, not all waves can propagate as if there were no other waves around. For example, sea waves change their shape when approaching the shore. This is a typical aspect of the **nonlinear** behaviour of waves.

The above terminology is associated with the mathematical form of the equations describing wave processes. For small amplitudes, we can neglect terms containing squares, products, and higher powers of amplitudes in the equations which thus become linear. If, however, the amplitudes are not small, this simplification is impossible, and the equations are nonlinear.

The quantitative criterion for linear waves is described by the condition

$$\frac{2A}{\lambda} \ll 1.$$

For example, the height of sea waves away from the shore is many times smaller than the wavelength, i.e. linear waves are quite smooth waves.

The criterion for nonlinear waves has the form

$$\frac{2A}{\lambda} \sim 1.$$

The same sea waves become high and steep near the shore. The change in their shape gives rise to instability, and the wave is partially ruptured.

6.8.4. Solitons and Other Nonlinear Effects. Small-amplitude linear waves in a medium obey the usual laws of reflection, refraction, interference, and diffraction. For steep nonlinear waves which also interact with one another upon being superimposed, a number of specific phenomena is observed.

One such phenomenon is the **self-constriction of wave packets.** Simple periodic waves that are nearly sine but have slightly sharpened crests are found to be unstable. A uniform bunch of such waves cannot retain its regular order for a long time and is broken into individual wave packets with a smaller amplitude. This results in a surge of waves. Such an effect is observed under the conditions when the liquid layer over the solid surface has dimensions $d \sim \lambda$.

In thin liquid layers, when the solid underlayer surface is at a depth $d \sim \lambda/10$, a nonlinear effect associated with the formation of a solitary wave, or a **soliton**, may be observed. Such a solitary wave has the form of a nonsinusoidal symmetric hump vanishing on either side without forming any troughs on the surface.

A soliton is a localized wave packet formed by surface (gravitational) waves. This is a self-sustained wave packet in which a balance is struck between two opposite tendencies: the wave packet tends to spread out on account of dispersion and at the same time experiences self-constriction due to nonlinearity. When two solitons meet, they pass through each other without destruction.

Seismic sea waves are sometimes formed as solitons. The famous floods in Leningrad are also probably due to solitons. The red spot in Jupiter's atmosphere is also assumed to be a long-lived soliton.

The concept of a soliton has assumed a general physical significance in recent years. In the arduous subjects concerning nonlinear physics, the concept of a soliton is as important as that of an oscillator in linear physics.

6.8.5. Highly Perturbed Media. So far we have considered only regular waves. However, one frequently encounters a situation, like a storm in a sea, in which a real wave is a random superposition of elementary waves of different wavelengths, amplitudes, phases, and directions. According to the probability laws, there is a rare possibility in such cases of the formation of exceptionally high waves, called the ninth waves.

In order to understand the physical meaning of this phenomenon, we must first find the regularities in random phenomena on the basis of the statistical observations and generalizations. The myriads of molecules in gases are moving

randomly in different directions at different velocities. However, the mean velocity of motion of molecules has quite a definite value and is linked unambiguously with parameters like the gas temperature. Statistical analysis of the results of observation of stormy waves shows in their hebaviour quite definite regularities whose nature, however, is not clearly known. As a matter of fact, the motion of a wave is collective in contrast to the individual motion of a molecule. The formation, evolution, and propagation of a wave may occur in many different forms.

The collective effects acquire a considerable significance in objects like a plasma. In view of the long-range Coulomb interaction between the ions and electrons constituting a plasma, all types of waves (electromagnetic, hydrodynamic, etc.) can easily be excited in it. For a highly excited plasma (called **turbulent plasma**), all the arguments concerning stormy waves are applicable. The turbulent plasma is the form of matter of stars like the Sun.

6.8.6. Oscillations of a Charged Drop and the Fission of Heavy Nuclei. Finally, let us study the behaviour of liquid drops. As the radius R of a drop increases, its volume V increases as R^3, while its area increases as R^2. When two drops merge into one, the volume of the resulting drop is equal to the sum of the volumes of the drops (this is so because the liquid is practically incompressible). However, the surface area of the resulting drop is smaller than the sum of the surface areas of the two drops. Hence the free surface energy decreases, and the process of merging of two drops becomes thermodynamically advantageous.

Suppose that a drop is electrically charged. Let us consider the condition of its stability against decomposition into smaller drops. For a drop carrying a charge Q, the force of rupture of the drop is $F_E \approx Q^2/R^2$, while the rupturing pressure $p_E \approx F_E/(4\pi R^2) = Q^2/(4\pi R^4)$. On the other hand, the surface tension creates a pressure $p_{s.t} = 2\alpha/R$ contracting the drop (here, α is the surface tension). Consequently, the stability boundary of the drop is determined by the relation $p_E \approx p_{s.t}$, or

$$\frac{Q^2}{4\pi R^4} \approx \frac{2\alpha}{R}. \qquad (*)$$

Let us estimate the charge carried by a water drop of radius $R = 10^{-3}$ m, which loses stability and dissociates into smaller particles. Substituting the value $\alpha \simeq 7 \times 10^{-2}$ N/m for water into Eq. (*), we obtain

$$Q \approx \sqrt{8\pi\alpha R^3}.$$

This formula is in the CGS units. Substituting the appropriate numerical values, we obtain $Q \simeq 10^{10}e$, where e is the electron charge.

In order to determine whether this is a small charge or not, we compare the charge of the drop with the number N of molecules in it. For this purpose, we write

$$N = \frac{\rho V}{m_{\substack{\text{water}\\\text{molecule}}}}, \quad N = \frac{10^3 \times \frac{4}{3}\pi (10^{-3})^3}{18 \times 1.6 \times 10^{-27}} \sim 10^{20}.$$

Hence the fraction of the uncompensated charge in a drop is

$$\frac{Q}{N} \sim \frac{10^{10}}{10^{20}} \sim 10^{-10}.$$

Let us now consider the fission of heavy nuclei. The substance forming a heavy nucleus can be treated as a nucleonic fluid, and the nucleus itself as a drop of such a fluid. For example, the nuclei of the fermium isotope $^{256}_{100}$Fm undergo a spontaneous fission with a half-life of 3.2 h. This means that after more than three hours, about half of a given quantity of Fm will be left. Thus, the fermium nuclei are practically at the stability boundary.

The fermium nucleus has a radius $R \simeq 10^{-14}$ m, and an electric charge $Q = 100e$. With the help of Eq. (*), we can find the order of magnitude for the surface tension of the nuclear matter:

$$\alpha \approx \frac{1}{8\pi} \frac{Q^2}{R^3}.$$

This formula is written in the CGS units. Converting the result into the SI units, we obtain

$$\alpha \sim 10^{17} \text{ N/m}.$$

For the sake of comparison, we give below the values of for some other substances:

Liquid helium	2.4×10^{-4}
Kerosene	2.4×10^{-2}
Mercury	0.55
Molten platinum	1.82
Water	7.2×10^{-2}

Let us now estimate the period of nuclear oscillations. For this purpose, we resort to dimensional analysis. Suppose that the equilibrium shape of a nucleonic drop is spherical. Then an oblate and a prolate ellipsoid can be regarded as distorted forms of this spherical drop. Suppose that the natural oscillations of an excited nucleus are caused only by surface tension.[2] Hence the oscillation period T can depend only on the surface tension α, the density ρ, and the size R of the drop.

After this, we follow the standard procedure (see Appendix II). We form a unique combination of the required quantity and the above-mentioned quantities:

$$T \sim \rho^x R^y \alpha^z,$$

where x, y, and z are unknown numbers. In order to find these numbers, we write the dimensions of the parameters appearing in this relation:

$$\mathrm{s} \sim (\mathrm{kg}\cdot\mathrm{m}^{-3})^x \cdot (\mathrm{m})^y \cdot (\mathrm{kg}\cdot\mathrm{s}^{-2})^z = \mathrm{kg}^{x+z}\cdot\mathrm{m}^{-3x+y}\cdot\mathrm{s}^{-2z}.$$

Comparing the exponents of terms on the left- and right-hand sides, we obtain

$$\left.\begin{aligned} x+z&=0,\\ -3x+y&=0,\\ -2z&=1 \end{aligned}\right\} \rightarrow x=\frac{1}{2}, \quad y=\frac{3}{2}, \quad z=-\frac{1}{2}.$$

Consequently,

$$T \sim \left(\frac{\rho}{\alpha}\right)^{1/2} R^{3/2}. \qquad (**)$$

For a nucleonic drop, i.e. a nucleus, the quantities appearing in Eq.(**) have the following values: $\rho \sim 10^{17}$ kg/m^3,

[2] For the sake of simplicity, we neglect the charge of the nucleus. Exact calculations show that this does not change the order of magnitude of the period of nuclear oscillations.

$R \simeq 10^{-14}$ m, and $\alpha \sim 10^{17}$ N/m. This gives the nuclear oscillation period $T \sim 10^{-21}$ s. Accordingly, the oscillation frequency is

$$\omega = \frac{2\pi}{T} \sim \frac{1}{T} \sim 10^{21} \text{ s}^{-1}.$$

6.9. Macroscopic Motion of Compressible Media

6.9.1. Generalized Form of the Bernoulli Equation. So far, we have considered the motion of fluids with velocities much smaller than the velocity of sound propagation in a medium. The description of such subsonic flows is based only on the laws of classical mechanics. However, if the velocities are close to the velocity of sound (this is the case of gas flows), we must take into consideration the compressibility of the medium (see below). In this case, one cannot formulate the equations of motion merely on the basis of the laws of mechanics. We must also take into consideration the laws of thermodynamics in this case.

In mechanics, the Bernoulli equation had the form

$$\frac{v^2}{2} + \frac{p}{\rho} = \text{const.} \tag{6.9.1}$$

In order to take into account the internal thermal energy, we add the internal energy ε per unit mass to the left-hand side of this equation. This gives

$$\frac{v^2}{2} + \frac{p}{\rho} + \varepsilon = \text{const.} \tag{6.9.2}$$

This equation can be written in a somewhat different form by introducing the specific enthalpy per unit mass:

$$h = \varepsilon + \frac{p}{\rho}, \tag{6.9.3}$$

where $h = H/m$ (see Eq. (5.8.14)). In this case, Eq. (6.9.2) assumes the form

$$\left(\frac{v^2}{2} + h\right) = \text{const.} \tag{6.9.4}$$

This equation is called the **generalized Bernoulli equation**.

6.9.2. Compressibility Criterion for a Medium and the Velocity of Sound. Let us consider the problem of a fluid flowing onto a solid body. Let h_0 be the specific enthalpy of a liquid particle (we consider a field tube) at a considerable distance ahead of the body, and let v_0 be its velocity. We choose a field tube such that the liquid particle flowing onto the body completely loses its velocity ($v = 0$); we denote the specific enthalpy of such a particle by h. Using Eq. (6.9.4), we obtain the following relation describing the flow:

$$\frac{v_0^2}{2} + h_0 = h. \tag{6.9.5}$$

Suppose that this steady-state flow is adiabatic (isentropic). In this case, $ds = 0$, where s is the entropy of the unit mass of the gas. From the thermodynamic formula

$$dh = T\,ds + \frac{1}{\rho}\,dp$$

we obtain the following relation for such flows:

$$dh = \frac{1}{\rho}\,dp. \tag{6.9.6}$$

When applied to the problem under consideration, this relation acquires the form

$$h - h_0 = \frac{1}{\rho_0}(p - p_0) \equiv \frac{\Delta p}{\rho_0} = \left(\frac{dp}{d\rho}\right)_s \frac{\Delta\rho}{\rho_0}.$$

Here, we assume that

$$p - p_0 \equiv \Delta p \ll p_0$$

in the flow of the liquid or gas. Consequently, (6.9.5) can be written as

$$\frac{v_0^2}{2} = \left(\frac{dp}{d\rho}\right)_s \frac{\Delta\rho}{\rho_0}. \tag{6.9.7}$$

It can be seen from this relation that $dp/d\rho$ has the dimensions of the squared velocity and is associated with the **compressibility** of the substance:

$$\frac{d\rho}{dp} = \left(\frac{dp}{d\rho}\right)^{-1}, \tag{*}$$

i.e. to the variation of its density with pressure. Obviously, compressibility depends on the magnitude of atomic (or

molecular) interactions. These interactions are quite strong in liquids, and any perceptible change in the separation between atoms (or molecules) requires a considerable external force. For example, to change the density of water just by 5%, an external pressure of 10^3 atm is required, while for a similar variation in the density of mercury, a pressure of 10^4 atm is needed. For a corresponding change in the density of air, a pressure of just 1.05 atm is sufficient.

Any slight change in the pressure over a small region of a medium resulting in a change in its density does not remain localized. On account of molecular interactions, this perturbation propagates over the medium in the form of alternating waves of compression and rarefaction. These are **sound** (or **acoustic**) **waves**, and their velocity of propagation is defined by the above-mentioned relation

$$u^2 = \left(\frac{dp}{d\rho}\right)_s. \tag{6.9.8}$$

Consequently, formula (6.9.7) becomes

$$\frac{\Delta\rho}{\rho_0} = \frac{1}{2}\frac{v_0^2}{u^2} \ll 1.$$

Hence, for a fluid moving with a velocity much smaller than the velocity of propagation of sound in this medium, we can neglect its compressibility. The change in the density of the medium becomes significant only at sonic and supersonic velocities of moving bodies in the medium.

The following are the values of sound velocity in different media. This velocity is 330 m/s in air under normal conditions, 1450 m/s in water, 6 km/s in solid iron, and (0.1-0.2)c in nuclear matter, c being the velocity of light.

6.9.3. Flow in a Tube with a Varying Cross Section. The flow conditions in this case can be formulated as follows: a gas moves isentropically in a tube with a slowly varying cross section, its flow velocity being defined as the mean value over its cross section, and the flow is assumed to be steady-state.

It is convenient to start the analysis from the mass flux conservation law:

$$\rho v S = \text{const}, \tag{6.9.9}$$

where S is the cross-sectional area of the channel. Since different quantities in this relation may change in different ways, we take the logarithm of this formula and differentiate it. This gives

$$\frac{d\rho}{\rho} + \frac{dv}{v} + \frac{dS}{S} = 0. \tag{6.9.10}$$

Since the flow is isentropic, the entropy $s = \text{const}$ for all parts of the flow. The first term in (6.9.10) can be written in the form

$$\frac{d\rho}{\rho} = \left(\frac{d\rho}{dp}\right)_s \frac{dp}{\rho} = \frac{dh}{u^2}. \tag{6.9.11}$$

Here, we have used formulas (6.9.6), (*), and (6.9.8). According to the generalized Bernoulli equation (6.9.4), we have

$$dh = -v\,dv. \tag{6.9.12}$$

Substituting (6.9.11) and (6.9.12) into (6.9.10), we arrive at the final expression

$$\frac{dv}{v}\left(1 - \frac{v^2}{u^2}\right) = -\frac{dS}{S}. \tag{6.9.13}$$

Let us analyze this relation.

1. For a subsonic flow ($v < u$), we have $(1 - v^2/u^2) > 0$:

(i) if a tube is convergent (i.e. $dS < 0$), we obtain $dv > 0$, which means that the flow velocity increases;

(ii) if a tube is divergent, i.e. has an increasing diameter, the subsonic flow is slowed down.

2. For a supersonic flow ($v > u$), we have $(1 - v^2/u^2) < 0$:

(i) in a converging tube ($dS < 0$), we obtain $dv < 0$, i.e. the flow velocity decreases;

(ii) in a diverging tube ($dS > 0$), we have $dv > 0$, i.e. the flow velocity increases.

The distinction between subsonic and supersonic flows of a gas can be seen quite clearly in the case of a channel with a varying cross section.

6.9.4. Laval Nozzle. A gas which is at rest in a combustion chamber is released to the atmosphere through a convergent-divergent nozzle (Fig. 6.16). According to Eq. (6.9.13),

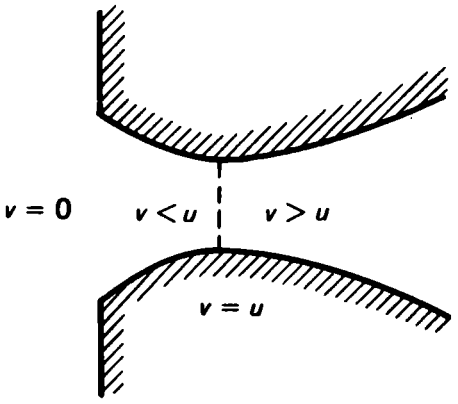


FIG. 6.16

the flow velocity may become equal to the velocity of sound ($v = u$) only at the narrowest part of the nozzle, i.e. for $dS = 0$. In the divergent part of the nozzle, the flow is supersonic.

These results can be explained physically as follows. A sufficiently high pressure is created in the combustion chamber. This pressure must be higher than the external pressure to such an extent that the gas rushing to the nozzle attains the velocity of sound at the narrowest part of the nozzle. Naturally, the pressure in the gas jet continuously falls in this process. In the divergent part of the nozzle, the gas is further accelerated at the expense of the internal energy, and its temperature steadily falls. Thus, a supersonic jet is formed owing to an increased pressure and the internal thermal energy of the jet itself.

This type of nozzle is used in a liquid-propellant rocket (LPR) engines and jet engines. The following are the parameters of an LPR engine operating on a chemical fuel. The pressure in the combustion chamber of a high-power LPR engine attains values of 80-100 atm, while the temperature reaches values of 3000-4700 K. The exhaust gases escaping from the nozzle have a velocity of 3-4.5 km/s.

6.10. Shock Waves

6.10.1. Propagation of Perturbations in a Compressible Gas Flow. The compressibility of a medium is responsible for the emergence of a new range of phenomena, viz. sound waves and **shock waves.** We have described more than once the propagation of sound waves as small perturbations of the medium. The following factor is worth noting here. Since the velocity of sound u is the rate at which the oscillation pattern is transported across the medium, it becomes clear that if sound propagates in a medium moving with a velocity v, this drift velocity of the medium is superimposed on the velocity of sound. This is the basic difference be-

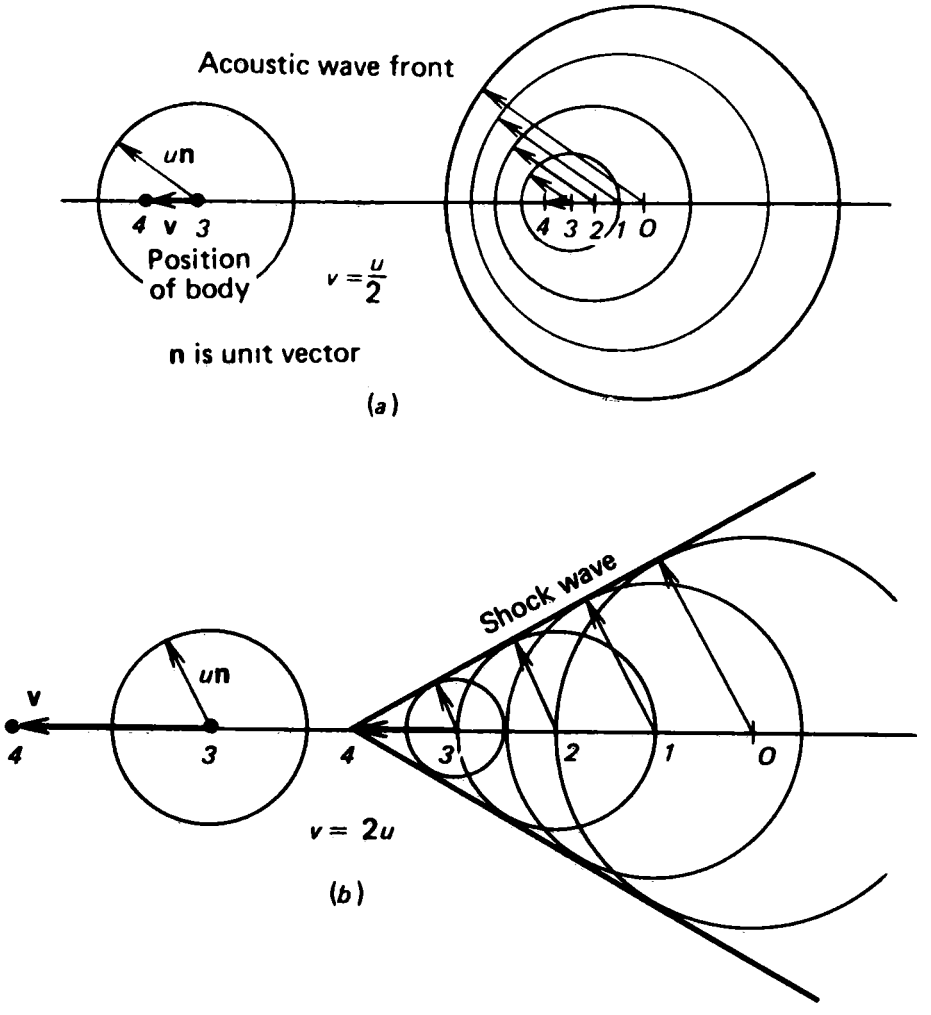


FIG. 6.17

tween the sound waves and electromagnetic waves since the velocity of the latter type of waves is independent of the choice of reference frame.

For a subsonic flow past a body, there is an all-round change in the space surrounding the body. In this case, the perturbations arising owing to the flow past the body propagate in all directions, as well as in front of the body. These perturbations as if "prepare" the flow for a contact with the body. Figure 6.17*a* shows the successive positions at instants 0, 1, 2, 3, 4 of a body moving with a subsonic velocity, as well as the cross sections of spherical fronts of sound waves emitted at various instants of time. Each circle shows the

position of the front at instant 4 for the wave emitted at instant 0, 1, 2, and 3 respectively.

For a supersonic flow past a body, perturbations cannot move in front of the body, and the flow "blindly" encounters the obstacle. In this case, perturbations exist only in the wake of the body. These perturbations are superimposed on one another and form a conical surface enveloping them with much higher values of density, pressure, and temperature. This is called the shock wave (Fig. 6.17*b*).

The difference in the behaviours of supersonic and subsonic gas flows was demonstrated above by considering the example of flows in tubes with a varying cross section. Here, we point out the basic aspect of a supersonic flow, i.e. the existence of surfaces on which all parameters describing the flow experience an abrupt increase in their values. These discontinuity surfaces separate the perturbed part of the flow from the unperturbed part. The perturbed flow is confined to a region bounded by the conical shock surface.

Before taking up a detailed investigation of the flow properties at a shock-wave discontinuity, let us consider a few quantitative estimates concerning sonic perturbations. We shall show that *the velocity of propagation of sound in a gas is nearly of the same order of magnitude as the mean thermal velocity of its molecules.* For this purpose, we use the equation of state for an ideal gas

$$pV \equiv \frac{p}{\rho} = NkT,$$

where $V = 1/\rho$ is the specific volume of the gas in the present case. Since

$$\left(\frac{dp}{d\rho}\right)_s \sim \frac{p}{\rho},$$

the velocity of sound is given, in accordance with Eq. (6.9.8), by

$$u \sim \sqrt{NkT}. \qquad (6.10.1)$$

This result shows that *the velocity of sound in a gas depends only on its temperature, and not on its pressure or density.*

Next, we make use of the relation

$$m\langle v^2\rangle \sim kT,$$

where m is the mass of a molecule, and $\langle v^2 \rangle$ is its mean square velocity. We can then write (6.10.1) in the form

$$u \sim \sqrt{Nm \langle v^2 \rangle}.$$

Since, by stipulation, we are considering a unit mass of the gas, $Nm = 1$. Consequently, the velocity of sound is given by

$$u \sim \langle v \rangle.$$

Following this derivation, let us verify the important fact that *the sonic or supersonic flow of a gas can be assumed to be nonviscous.* Indeed, the expression for the kinematic viscosity has the form

$$\nu \sim \langle v \rangle l \sim ul,$$

where l is the mean free path of molecules, and $\langle v \rangle$ is the mean thermal velocity of the molecules which was shown above to have the same order of magnitude as the velocity of sound u. Since the velocity v of macroscopic motion is also comparable with the velocity of sound, the Reynolds number

$$\mathrm{Re} = \frac{vL}{\nu} \sim \frac{uL}{ul} \sim \frac{L}{l} \tag{6.10.2}$$

assumes very large numerical values (L is the characteristic size, in the present case, the size of the body past which the gas flows). Large values of the Reynolds number correspond to a low viscosity.

6.10.2. General Relations for a Shock Wave. How can we quantitatively describe shock waves? Although the parameters ($\mathbf{v}$, p, ρ) of a gas undergo a jump as we go over through the "discontinuity surface", the flows corresponding to them must be continuous. This follows from the mass, momentum, and energy conservation laws.

Let us consider an element of the surface of a shock wave perpendicular to the flow and attach to this element a reference frame with the X-axis directed along the normal to the surface (Fig. 6.18). Then, in view of the conservation of matter, the mass flux density must be the same on both sides of the discontinuity surface, i.e.

$$\rho_1 v_{1x} = \rho_2 v_{2x}. \tag{6.10.3}$$

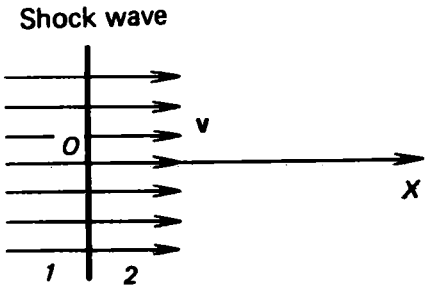


FIG. 6.18

Let us consider the continuity condition for the momentum flux at the discontinuity surface. The momentum of a unit volume of the gas is ρv. It would appear that the momentum flux density in the direction of the gas flow should be

$$(\rho v_x)\, v_x,$$

while the condition at the discontinuity surface could be written as the equality of these expressions on both sides of the shock wave. However, such a relation would not be complete since if the gas is at rest, the pressures exerted by neighbouring regions of the gas on each other must be equal. As a result, we obtain

$$p_1 + \rho_1 v_{1x}^2 = p_2 + \rho_2 v_{2x}^2. \tag{6.10.4}$$

On the basis of similar considerations, we can write the expression for the energy flux. The total energy of a unit volume of the gas has the form

$$\rho\left(\frac{v^2}{2} + \varepsilon\right).$$

Accordingly, for the energy flux density, we obtain

$$\rho v_x\left(\frac{v^2}{2} + h\right).$$

Here, we used relation (6.9.3). The continuity condition for the energy flux at the discontinuity surface has the form

$$\rho_1 v_{1x}\left(\frac{v_1^2}{2} + h_1\right) = \rho_2 v_{2x}\left(\frac{v_2^2}{2} + h_2\right). \tag{6.10.5}$$

It should be noted that in view of condition (6.10.3), this condition can be simplified. Taking into account the fact that with our choice of orientation of the reference frame $v_x = v$, the above conditions at the discontinuity surface can be written in the form

$$\rho_1 v_1 = \rho_2 v_2 \equiv j, \tag{6.10.6}$$

$$p_1 + \rho_1 v_1^2 = p_2 + \rho_2 v_2^2, \tag{6.10.7}$$

$$h_1 + \frac{v_1^2}{2} = h_2 + \frac{v_2^2}{2}. \tag{6.10.8}$$

A shock wave in a gas has a finite width and is a transient nonequilibrium layer separating two equilibrium regions of a medium with different values of the thermodynamic parameters characterizing them. Let us derive the relations which connect the thermodynamic parameters on both sides of the shock-wave discontinuity. We shall proceed from the general conditions (6.10.6)-(6.10.8). Let us introduce specific volumes $V_1 = 1/\rho_1$ and $V_2 = 1/\rho_2$ of the gas on both sides of the discontinuity. From (6.10.6), we obtain

$$v_1 = jV_1, \quad v_2 = jV_2. \tag{6.10.9}$$

Substituting these expressions into (6.10.7), we get

$$p_1 + j^2V_1 = p_2 + j^2V_2, \tag{6.10.10}$$

or

$$j^2 = \frac{p_2 - p_1}{V_1 - V_2}. \tag{6.10.11}$$

Since j^2 is a positive quantity, we must simultaneously have[3]

$$p_2 > p_1, \quad V_1 > V_2. \tag{6.10.12}$$

Thus, *the pressure of the gas behind the shock wave is higher than in front of it. The same refers to the gas density* ($\rho_2 > \rho_1$).

Using Eq. (6.10.11), we can write the following expression for the velocity jump $v_1 - v_2$ of the gas on both sides of the shock wave which, according to Eq. (6.10.9), is equal to $j(V_1 - V_2)$:

$$v_1 - v_2 = \sqrt{(p_2 - p_1)(V_1 - V_2)}. \tag{6.10.13}$$

Inequalities (6.10.12) lead to

$$v_1 > v_2. \tag{6.10.14}$$

Thus, *the velocity of the flow drops behind the shock wave.*

Let us now consider condition (6.10.8) which we write in the form

$$h_1 + \frac{j^2V_1^2}{2} = h_2 + \frac{j^2V_2^2}{2}. \tag{6.10.15}$$

[3] A more detailed analysis shows that the inverse inequalities cannot be valid in this case.

Here, we have used equalities (6.10.9). Substituting (6.10.11) into (6.10.15), we obtain

$$h_1 - h_2 + \frac{1}{2}(V_1 + V_2)(p_2 - p_1) = 0. \tag{6.10.16}$$

The value of enthalpy behind the shock wave is higher than in front of it ($h_2 > h_1$).

In Eq. (6.10.16), we go over from the enthalpy

$$h = \varepsilon + pV$$

to the internal energy ε. This gives

$$\varepsilon_1 - \varepsilon_2 + \frac{1}{2}(V_1 - V_2)(p_1 + p_2) = 0. \tag{6.10.17}$$

The internal energy of the gas increases behind the shock wave ($\varepsilon_2 > \varepsilon_1$).

Relations (6.10.16) and (6.10.17) are very important for the theory of shock waves since they determine the relation between the thermodynamic quantities on both sides of the shock-wave surface.

As to the change in entropy, it can only increase for a gas passing through the shock layer (in view of the energy degradation law), i.e.

$$s_2 > s_1. \tag{6.10.18}$$

At first glance, this inequality contradicts to what was said at the end of the previous section (see Sec. 6.10.1). Namely, the flow of a gas with sonic and supersonic velocities can be regarded as the flow of an "ideal liquid", i.e. a medium with zero viscosity and thermal conductivity. It is well known that the entropy remains unchanged for such a flow. Therefore, shock-wave discontinuities can be treated as a mechanism of increasing entropy of an "ideal liquid". In other words, the properties of viscosity and thermal conductivity of the medium will be manifested over the width of a shock wave. Since the viscosity is significant for flows with small values of the Reynolds number, it follows from Eq. (6.10.2) that *the width L of a shock wave can be only of the order of the mean free path l of molecules.*

This result does not contradict the assumption that in the

gas dynamics of a continuum the shock front is a geometrical surface. As a matter of fact, gas dynamics neglects the molecular structure of the medium, assuming that $l \to 0$. As regards the structure and processes occurring over the width of a shock wave, they can be analyzed completely and consistently by the methods of physicochemical statistical kinetics.

Concluding this section, it should be emphasized that while deriving the relations and inequalities concerning the behaviour of thermodynamic parameters for the transition through a shock wave, we have never used the equation of state for a substance. This means that the relations and inequalities obtained by us are of general nature and can be applied to shock waves propagating in any medium: solid, liquid or gaseous.

6.10.3. Shock Waves in an Ideal Gas. Here we shall specify the medium: we consider the simplest model of an ideal molecular gas. It should be recalled that the term "ideal gas" in statistical thermodynamics implies that we have a gas of weakly interacting particles. The equilibrium properties of such a gas were considered in Chap. 5. For a theoretical analysis of the behaviour of thermodynamic parameters of the gas at a shock-wave discontinuity using the general relations derived in Sec. 6.10.2, we shall require the equation of state for an ideal gas and the expressions for its internal energy and enthalpy.

First of all, let us derive the equations for the internal energy and enthalpy of an ideal gas in the form convenient for our purposes. It follows wrom Eq. (5.8.12) that to within an additive constant, the internal energy of an ideal gas with the temperature-independent thermal conductivity has the form

$$\varepsilon = c_V T. \tag{6.10.19}$$

Expressing the temperature from the equation of state (5.8.4) and using expression (5.8.18) for the gas constant, we can write (6.10.19) in the form

$$\varepsilon = \frac{c_V}{c_p - c_V} pV = \frac{pV}{\gamma - 1}, \tag{6.10.20}$$

where $\gamma = c_p/c_V$. Similarly, from Eq. (5.8.15) we obtain the expression for the enthalpy of an ideal gas:

$$h = c_p T = \frac{\gamma}{\gamma - 1} pV. \tag{6.10.21}$$

We now have everything to calculate the discontinuities of the basic parameters at the discontinuity surface in the ideal gas. The analysis will be carried out so that the ratios of similar quantities at the shock front be expressed in terms of the gas pressure p_1 in front of the shock wave and the gas pressure p_2 behind the shock wave (an expression may also contain the initial density $\rho_1 = 1/V_1$ of the gas). The ratio p_2/p_1 specifies the intensity of a shock wave, and hence it can be treated as an arbitrarily varying parameter.

We can now proceed with the calculations. Substituting (6.10.21) into the general relation (6.10.16), we obtain[4]

$$\frac{\gamma}{\gamma - 1} (p_1 V_1 - p_2 V_2) + \frac{1}{2} (V_1 + V_2)(p_2 - p_1) = 0,$$

or

$$\frac{V_2}{V_1} = \frac{(\gamma + 1) p_1 + (\gamma - 1) p_2}{(\gamma - 1) p_1 + (\gamma + 1) p_2}. \tag{6.10.22}$$

Using the relation

$$\frac{p_1 V_1}{T_1} = \frac{p_2 V_2}{T_2}, \tag{6.10.23}$$

which follows from the equation of state for an ideal gas, we obtain the following expression for the ratio of temperatures on both sides of the shock wave:

$$\frac{T_2}{T_1} = \frac{p_2}{p_1} \left(\frac{V_2}{V_1} \right) = \frac{p_2}{p_1} \left[\frac{(\gamma + 1) p_1 + (\gamma - 1) p_2}{(\gamma - 1) p_1 + (\gamma + 1) p_2} \right]. \tag{6.10.24}$$

Here, we have used expression (6.10.22) for V_2/V_1.

The formula for the mass flux density j can be obtained by substituting $V_1 - V_2$ from (6.10.22) into (6.10.11). This gives

$$j^2 = \frac{(\gamma - 1) p_1 + (\gamma + 1) p_2}{2V_1}. \tag{6.10.25}$$

[4] It should be recalled that the subscripts 1 and 2 indicate the values of a quantity in front of the shock wave and behind it respectively.

For the velocities of a gas flow relative to the discontinuity surface, we obtain

$$v_1^2 = \frac{V_1}{2}[(\gamma-1)p_1+(\gamma+1)p_2],$$
$$v_2^2 = \frac{V_1}{2}\frac{[(\gamma+1)p_1+(\gamma-1)p_2]^2}{(\gamma-1)p_1+(\gamma+1)p_2}. \quad (6.10.26)$$

In order to derive these relations, we have used Eqs. (6.10.9) and (6.10.25). In deriving the second relation of (6.10.26), the expression for V_2^2 was taken from (6.10.22). It should be noted that formulas (6.10.26) describe not only the velocity of the gas flow relative to the shock-wave surface but also the velocity of propagation of the shock wave relative to the gas in front of the wave and behind it.

Let us consider high-intensity shock waves in which $p_2 \gg p_1$. Then it follows from Eq. (6.10.22) that

$$\frac{V_2}{V_1} = \frac{\rho_1}{\rho_2} = \frac{\gamma-1}{\gamma+1}. \quad (6.10.27)$$

This relation shows that the ratio of densities tends to a constant limit. Thus, according to (5.8.20), for a monatomic gas we have

$$\gamma = \frac{c_p}{c_V} = \frac{5R/2}{3R/2} = \frac{5}{3},$$

i.e.

$$\frac{\rho_1}{\rho_2} = \frac{\gamma-1}{\gamma+1} = \frac{(5/3)-1}{(5/3)+1} = \frac{1}{4},$$

or

$$\rho_2 = 4\rho_1.$$

According to Eq. (5.8.21), for a diatomic gas we obtain

$$\gamma = \frac{7}{5}, \quad \rho_2 = 6\rho_1.$$

In the limiting case of high-intensity shock waves, from Eq. (6.10.24) we obtain the following relation for the temperature ratio:

$$\frac{T_2}{T_1} = \frac{(\gamma-1)}{(\gamma+1)}\frac{p_2}{p_1}. \quad (6.10.28)$$

Therefore, the ratio T_2/T_1 increases unlimitedly with p_2/p_1. In other words, *the temperature jump in a shock wave, like the pressure jump, can be as large as desired.*

Finally, for the velocity of propagation of high-intensity shock waves we obtain from Eq. (6.10.26)

$$v_1 = \sqrt{\frac{\gamma+1}{2} p_2 V_1}, \quad v_2 = \sqrt{\frac{(\gamma-1)^2}{2(\gamma+1)} p_2 V_1}. \tag{6.10.29}$$

These velocities increase in proportion to the square root of the pressure p_2.

Shock waves can emerge as a result of an explosion. *The velocity of propagation of a shock wave is higher than the velocity of sound.* It follows, for example, from the first relation in (6.10.26) which is transformed to

$$v_1 \to \sqrt{\gamma \frac{p_1}{\rho_1}} = u, \tag{6.10.30}$$

as $p_2 \to p_1$. This means that *low-intensity shock waves propagate with a velocity close to the velocity of sound.*

The pressure and the energy density in a strong shock wave are much higher than in an unperturbed gas. By way of an example, let us consider a wave with a pressure of 100 atm in air. This wave compresses air almost by a factor of eight and propagates with a velocity exceeding 3 km/s, i.e. ten times as high as the velocity of sound. The temperature in it reaches 3500 °C. At this temperature, a noticeable fraction of air molecules have already dissociated, i.e. decomposed into atoms. For a much larger remaining part of the molecules, not only translational and rotational but also vibrational degrees of freedom become significant at this temperature. The limiting compression in this case is eightfold instead of the six-fold compression in the absence of the vibrational states.

A shock wave with a pressure of 1000 atm heats the air to 14 000 °C. At such a temperature, electron atomic states "come into play" (see Sec. 4.10). Transitions of atoms from the ground state to excited states are observed. The inverse transitions are accompanied by the emission in the visible region of the spectrum. The limiting compression ratio in this case is equal to ten. The heated air becomes opaque for its own radiation. The front surface of the strong shock wave in a gas emits like a red-hot solid at the same temperature.

The energy emitted by a square centimetre of the surface of such a wave per second is 36 times as high as the energy emitted by the same surface element of the Sun. It should be recalled that the temperature of the emitting layer of the Sun is 5700 °C.

During a nuclear explosion, brightly glowing shock waves like balls of fire are formed. If a ball of fire at a temperature of 14 000 °C is viewed at an angle which is 5.3 times larger than that of the solar disc, it glows brighter than a thousand of suns ($36 \times 5.3^2 > 1000$). However, such a bright glow lasts for a short time (less than a hundredth of a second).

6.10.4. The Problem on a High-Intensity Explosion in the Atmosphere. The propagation of strong shock waves emitted from a "point" source in the atmosphere (say, during a nuclear explosion) obeys a certain similitude law (see Sec. 6.6.1). It should be interpreted as follows. If the radius of a spherical shock wave formed during an explosion increases twofold or three-fold, the distribution of all gas-dynamic parameters in the wave remains the same if the corresponding quantities at the wave front are taken as units of measurement.

Let us find out why *high-intensity spherical shock waves are self-similar*. For this purpose, let us consider a local explosion accompanied by the liberation of a large amount of energy E. The gas itself (atmosphere) in which the high-intensity shock wave propagates will be assumed to be an ideal gas of density ρ. In this problem, we can neglect in advance the initial energy and pressure of the air. The range R of distances from the "point" of explosion (origin) is such that the shock waves remain strong enough.

We must derive the law of propagation of a shock wave, i.e. the dependence $R(t)$ for given parameters E and ρ. Let us try to compile a dimensionless combination of these quantities. First of all, since energy is expressed in joules ($1\ \text{J} = 1\ \text{kg}\cdot\text{m}^2/\text{s}^2$), and density in kilograms per cubic metre, the required relation between E and ρ should include these quantities in such a way that kilograms be cancelled out. It is natural to take their ratio

$$\left[\frac{E}{\rho}\right] = 1\ \text{m}^5/\text{s}^2.$$

It only remains for us to multiply this ratio by t^2/R^5 to obtain the required dimensionless combination

$$\left(\frac{Et^2}{\rho R^5}\right) \sim 1. \qquad (*)$$

It is remarkable that this combination is unique, and this ensures the fulfilment of the similitude law for high-intensity spherical shock waves propagating from a point source. It follows from Eq. (*) that

$$R \sim \left(\frac{E}{\rho}\right)^{1/5} t^{2/5}. \qquad (6.10.31)$$

For the velocity of propagation of the shock wave, we obtain

$$v = \frac{dR}{dt} \sim \left(\frac{E}{\rho}\right)^{1/5} t^{-3/5}. \qquad (6.10.32)$$

Equation (6.10.29) leads to the following expression for pressure in the high-intensity shock wave:

$$p = \frac{2}{\gamma+1} \rho v^2$$

(here, we omitted all subscripts). Substituting the expression for velocity obtained above, we get

$$p \sim (E^{2/5}\rho^{3/5})\, t^{-6/5}. \qquad (6.10.33)$$

According to Eq. (6.10.27), the density itself does not change with time.

The direct experimental verification of relations (6.10.31)-(6.10.33) carried out during nuclear explosions confirmed their validity to a high degree of accuracy.

6.11. Hydrodynamic Cumulative Effects

In continua, the **cumulative effects** are observed, in which the volume energy density spontaneously increases. The computational methods show that the energy density could increase indefinitely. However, instability sets the limits to the cumulative effects.

Cumulative effects include the collapse of bubbles in liquids, converging shock waves (which form cumulative jets), and the emergence of shock waves on the surface of a star.

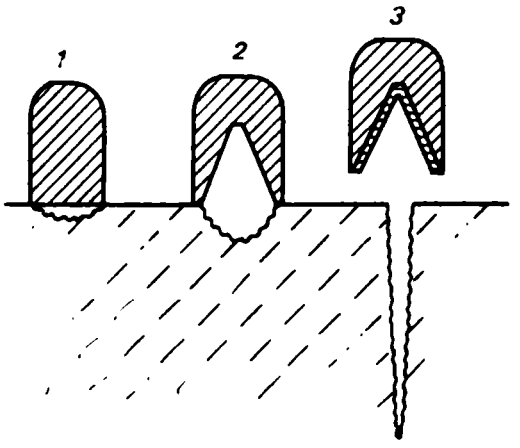


FIG. 6.19

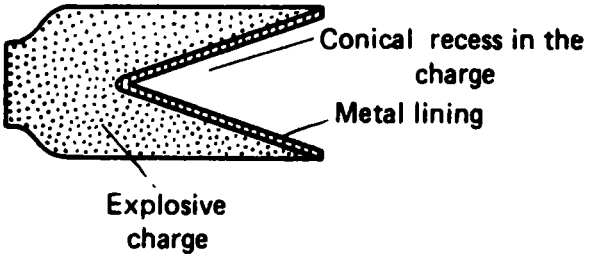


FIG. 6.20

6.11.1. Cumulative Jets. Let us consider a set of charges of an explosive having the same mass but different shapes (Fig. 6.19). Sample *1* is solid, while sample *2* has a conical recess. In sample *3*, such a recess is lined with a metal (Fig. 6.20). Let us place charges *1* and *2* on an armour slab, while charge *3*, at a certain distance from the slab. The results of explosions of the charges are shown in Fig. 6.19.

During the explosion of charge *3*, a thin metal jet having a velocity of the order of escape velocity ($\sim$ 10 km/s) is formed. The mechanism of its formation is quite simple: the explosion causes a compression of the metal cone, and the shock-type collapse of the lining of the cone transforms the metal into the liquid state. The energy is concentrated (cumulated) along the cone axis, and the cumulative liquid-metal jet is ejected in the direction of the axis.

How can we describe this phenomenon quantitatively? First of all, this requires a number of numerical estimates. Let the conical lining be made of iron. The velocity v acquired

by the plates of the lining during the explosion is about 1 km/s. In order to estimate the pressure emerging during the collapse of the plates, we shall use the formula for pressure in a high-intensity shock wave:

$$p = \rho_0 D v, \tag{6.11.1}$$

where ρ_0 is the initial density of the substance, D is the velocity of propagation of the shock front, and v is the mass velocity (the velocity of motion of the compressed substance). If we are interested only in the orders of magnitude of the quantities, we can write $D \sim u$, where u is the velocity of sound in the substance. Then $p \sim \rho_0 u v$.

For iron, $u = 5$ km/s and $\rho_0 = 8 \times 10^3$ kg/m^3. Therefore,

$$p \sim 8 \times 10^3 \times 5 \times 10^3 \times 10^3 = 4 \times 10^{10}\,\text{Pa} = 4 \times 10^5\,\text{atm}.$$

Let us briefly consider the method of obtaining formula (6.11.1). Conditions (6.10.6)-(6.10.8) at the shock discontinuity are written in the reference frame fixed to the wave front. Let us go over to a reference frame in which the substance in front of the wave is at rest. Then instead of the variables v_1 and v_2, we shall have D and $D - v$. For a high-intensity shock wave ($p_1 = 0$), expressions (6.10.6) and (6.10.7) can be written as

$$\rho_1 D = \rho_2 (D - v), \tag{6.11.2}$$

$$\rho_1 D^2 = p_2 + \rho_2 (D - v)^2. \tag{6.11.3}$$

Taking into account (6.11.2), we obtain from (6.11.3) the required relation (6.11.1) in which all the subscripts are omitted.

Earlier, we estimated the mechanical strength of ionic defect-free crystals and obtained the value $\sim 10^{10}$ N/m^2. For real iron samples, the value of strength is much lower. Thus, as a result of collapse, the iron plates will go over to the liquid state.

As regards the relative compression of the liquid metal, its value can be estimated from the relation

$$\frac{\Delta\rho}{\rho} \approx \frac{v}{D} \approx \frac{v}{u} = \frac{1}{5} = 0.2,$$

which follows from (6.11.2). Such a compression can be neglected, and the metal liquid formed can be regarded as incompressible.

The above analysis allows us to choose the approximation of an ideal incompressible liquid. The kinematics of motions during the formation of cumulative jets is described by the ordinary Bernoulli equation.

Let us take advantage of the idea about the **laboratory reference frame (lab-system)** and the **centre-of-mass reference frame (c-system)**. In order to clarify their difference, let us consider the decay of a particle. Let a parent particle move before the decay in the lab-system with a velocity **V**. In the c-system, this particle is at rest, i.e. its total momentum is zero. Denoting the velocities of the daughter particle in the lab- and c-systems by $\mathbf{v}_{lab}$ and $\mathbf{v}_c$ respectively, we can write the conventional relation

$$\mathbf{v}_{lab} = \mathbf{v}_c + \mathbf{V}.$$

The diagram of formation of a cumulative jet is shown in Fig. 6.21*a* for the lab-system. According to the diagram of velocity composition (Fig. 6.21*b*), the diagram of flows in the c-system fixed to point *O* has the form shown in Fig. 6.21*c*. The jets from sections *1-1*, meeting at point *O*, generate the flows to the right and to the left. As was mentioned above, all these steady-state flows are described by the Bernoulli equation. For example, for sections *1-1* and *2-2* we can write

$$p_1 + \frac{\rho v_{c1}^2}{2} = p_2 + \frac{\rho v_{c2}^2}{2}.$$

Here, we can assume that for sections *1-1* and *2-2* far from the collapse region, $p_1 = p_2 = 0$, i.e. $v_{c1} = v_{c2}$. The same refers to sections *1-1* and *3-3*. In other words, in the c-system, the velocities of all the three jets are equal.[5]

In the lab-system, the velocity of the cumulative jet is

$$v_{c.j} = v_c + V,$$

where V is the velocity of point *O*. According to Fig. 6.21*b*, we have $V = v_{lab}/\sin\alpha$ and $v_c = v_{lab}/\tan\alpha$. As a result, we obtain

$$v_{c.j} = \frac{v_{lab}}{\tan\alpha} + \frac{v_{lab}}{\sin\alpha} = v_{lab}\frac{1+\cos\alpha}{\sin\alpha}. \qquad (6.11.4)$$

[5] The symmetry of the problem allows us to confine ourselves to just three jets.

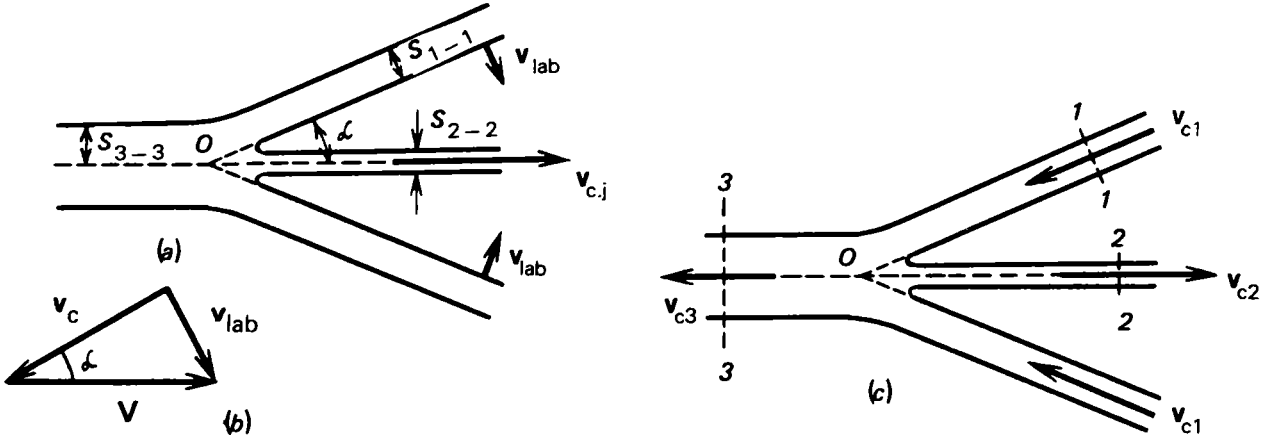

S_{1-1}
v_{lab}
S_{3-3}
O
α
S_{2-2}
$v_{c.j}$
v_{lab}
(a)
v_c
v_{lab}
α
V
(b)
1
v_{c1}
1
3
2
O
v_{c2}
v_{c3}
2
3
(c)
v_{c1}

FIG. 6.21

This relation shows that for small angles α of collapse of the plates, the velocity of a cumulative jet can be quite high ($v_{c.j} \to \infty$ as $\alpha \to 0$).

Let us derive the expressions for the cross-sectional area of a cumulative jet and the fraction of mass entrained by it. We proceed from the conditions of conservation of mass and of the horizontal momentum component. Then the mass flow rates for all the three cross sections in the c-system are related as follows:

$$\rho v_{c1} S_{1-1} = \rho v_{c2} S_{2-2} + \rho v_{c3} S_{3-3}.$$

For the horizontal component of the momentum per unit time, we obtain

$$-\rho S_{1-1} v_{c1}^2 \cos\alpha = \rho S_{2-2} v_{c2}^2 - \rho S_{3-3} v_{c3}^2.$$

The last two relations can be written in the form

$$S_{1-1} = S_{2-2} + S_{3-3},$$
$$-S_{1-1} \cos\alpha = S_{2-2} - S_{3-3}.$$

Summing them up, we obtain the fraction of mass carried away by the cumulative jet:

$$\frac{S_{2-2}}{S_{1-1}} = \frac{1-\cos\alpha}{2}. \qquad (6.11.5)$$

As $\alpha \to 0$, this fraction becomes as small as desired.

In spite of this circumstance, the cumulative jet carries away the major portion of the energy of the plates. Indeed, writing the ratio of the kinetic energies of the cumulative jet and of the regions of unit length of the plates, we obtain

$$\frac{E_{c.j}}{E_{lab}} = \frac{\rho S_{2\text{-}2} v_{c.j}^2}{\rho S_{1\text{-}1} v_{lab}^2} = \frac{1+\cos\alpha}{2}. \qquad (6.11.6)$$

Here, we have used Eqs. (6.11.4) and (6.11.5).

Relation (6.11.6) shows that the ratio $(E_{c.j}/E_{lab}) \to 1$ as $\alpha \to 0$. As follows from Eq. (6.11.4), the energy density in the cumulative jet increases indefinitely.

Thus, an analysis of the formation of cumulative jets in the framework of the ideal incompressible liquid approximation leads to an unlimited cumulation of energy. If the compressibility is taken into account, the cumulation becomes limited.

Cumulative jets find an application in physical experiments, for one, in high-pressure physics. A number of large-scale astronomical experiments can be explained with the help of the cumulative model. Cumulative projectiles have been used in military science for a long time. They are able to make narrow and deep shell holes in armour. A cumulative water jet produced by a "hydraulic gun" easily destroys hard rocks (like coal seams), which considerably simplifies their extraction.

6.11.2. Bubble Collapse in a Liquid. If an empty spherical cavity is formed in a liquid, it will be rapidly filled under the action of the pressure exerted by surrounding regions. Suppose that the motion of the liquid in this case is spherically symmetric. The liquid itself is inviscid and is characterized by the density ρ and the pressure p_0 at the boundary of a bubble. Then the initial total energy is

$$E = p_0 \cdot \frac{4}{3} (\pi R_0^3),$$

where R_0 is the initial radius of the bubble.

It is clear from the simple dimensional considerations that the pressure p and the velocity v of the boundary of the cavity vary during the collapse with the radius R according to the laws

$$p \sim \frac{E}{R^3}, \quad v^2 \sim \frac{E}{\rho} \frac{1}{R^3}.$$

As a result, as $R \to 0$ (focussing of the cavity), the pressure and the velocity of the boundary increase indefinitely. With increasing pressure, the concentration (cumulation) of the energy density is observed.

The problem on the collapse of an empty spherical cavity in a liquid is similar to the problem on a high-intensity explosion in the atmosphere (see Sec. 6.10.4). The process of collapse is the time inversion of the process of propagation of the shock front during a high-intensity explosion. Therefore, we can assume that the cavity radius varies with time t according to the law

$$R \sim \left(\frac{E}{\rho}\right)^{1/5} (-t)^{2/5},$$

while the velocity of the cavity boundary varies as

$$v \sim \left(\frac{E}{\rho}\right)^{1/5} (-t)^{-3/5}.$$

Here, for the time $t = 0$ we take the moment of focussing.

Analyzing the collapse of the empty cavity in an incompressible inviscid liquid, we come across the phenomenon of unlimited cumulation. Is this an accidental consequence of the idealization in the formulation of the problem? A consistent account of the viscosity of the medium leads to the conclusion that viscosity eliminates cumulation only in very small bubbles (for example, the size of such bubbles in water must be smaller than 0.8×10^{-6} m). Taking compressibility into consideration, we do not eliminate an unlimited cumulation. The limit on cumulation is undoubtedly set by the atomic theory since for the size of the bubbles comparable with atomic dimensions, the medium cannot be treated as continuous, and cumulation ceases.

The collapse of bubbles in a liquid is of a considerable practical interest since it is one of the reasons behind the rapid wear of the blades of propellers and hydraulic turbines operating under cavitation (the collapse of bubbles on a metal surface rapidly destroys it). Cavitation will be discussed below.

6.11.3. Converging Spherical and Cylindrical Shock Waves. Converging spherical and cylindrical shock waves are another example of an unlimited cumulation. Here, the density of matter is always limited during the focussing, while the velocity, pressure, and temperature tend to infinitely high values.

Converging shock waves produce huge temperature gradients near the centre, which leads to energy dissipation due to the heat conduction (and radiation) of the medium. However, calculations show that the latter circumstance cannot set a limit on cumulation. Other mechanisms should be found for its suppression.

6.11.4. The Role of Instabilities in Suppressing Cumulation. During cumulation, small perturbations may lead to the suppression of the process. A simple and visual example is a thin cylindrical envelope of an ideal liquid which slowly

rotates and rapidly converges to the axis. In this convergence, the envelope itself gradually becomes thicker (since the substance does not vanish in the process), and the rotation is intensified (owing to angular momentum conservation). The increase in the centrifugal force does not allow the envelope to reach the axis, which is followed by the divergence of the envelope. Therefore, an unlimited cumulation does not take place in this case, the reason behind that being a small perturbation in the form of a slow rotation, which gradually absorbs the entire energy of the main motion.

The problem on the liquid cylindrical envelope converging to the axis considered above has found an interesting practical application in producing superstrong magnetic fields. A copper envelope of the cylindrical shape (i.e. liner) is surrounded from outside by an explosive charge. The maximum possible longitudinal magnetic field is produced for a short time within the cylinder, and at the same moment the explosive is ignited. The shock effect of the explosive on the liner transforms it to the liquid state. Liquid copper has a high electric conductivity, and the magnetic field practically cannot penetrate the liner walls (this effect will be considered below). As a result of the axial compression of the liner, the magnetic field strength increases in the same proportion as the volume of the inner cavity of the liner decreases (practically, by two orders of magnitude). In this way, it has become possible to produce pulsed magnetic fields of strength 10^8-10^9 A/m.

6.11.5. Emergence of a Shock Wave on the Surface of a Star. The cumulative effects considered above are very important for astrophysical problems. When internal perturbations accompanied by an abrupt increase in pressure occur in the central region of a star, a shock wave formed propagates from the centre to the periphery and emerges on the surface. The propagation of the shock wave in a gas whose density drops to zero (near the surface) is accompanied by the concentration (cumulation) of energy. As a matter of fact, the energy is transferred to an infinitely decreasing mass of the substance so that the specific energy (viz. the energy per unit mass) grows unlimitedly.

There is a physical similitude between the cumulation during the propagation of a shock wave in a gas whose den-

sity drops to zero and the convergence of a shock wave to the centre. The difference lies only in the reasons behind the decrease in the mass to which energy is transferred. In the former case, the mass decreases with a decreasing gas density, and in the latter case, with a decreasing volume.

During the burst of a supernova, we observe the stripping of its envelope and apparently the formation of the flow of high-energy particles with $E \sim 10^{17}$ eV. These phenomena can be explained within the framework of the theory of energy cumulation considered above.

As regards "terrestrial applications", cumulative effects may play a significant role for creating extremal states of matter and electromagnetic field (attainment of high pressure, temperature, and field strength). For example, in the very important and complicated problem of controlled nuclear fusion, some trends are associated with energy cumulation. They include, in particular, pulsed systems for initiating nuclear fusion reactions by laser beams converging to the centre (on a deuterium-tritium target) or, in another version, converging relativistic electron beams. A rapid shock compression of a hydrogen plasma by an external magnetic field also belongs to such effects.

6.12. Cavitation in a Liquid

Microscopic bubbles with a size of about 10^{-9}-10^{-8} m are always present in a liquid. One of the reasons behind this effect is the natural cosmic corpuscular background of radiation. A decrease in pressure above the liquid may cause an intense evaporation in the bubbles and lead to their growth.[6] A gas (air) dissolved in the liquid also facilitates the bubble growth.

The growth of bubbles leads to their collapse. The duration of the entire process is of the order of 10^{-3} s. During a collapse, a shock wave is formed, leading to an abrupt increase in the temperature and pressure of the gas in the bubble. The estimates and measurements show that the temperature and pressure drop in the vicinity of a bubble may be 500-800 °C and 4000 atm respectively. The gas contained

---

[6] Compare with bubble chambers used for detecting high-energy particles.

in the bubble can emit weak visible radiation under the action of the shock wave.

When a liquid flows past a body with a high velocity (about 40-50 m/s), a pressure drop is observed in the regions of maximum velocity, which may cause local boiling of the liquid, accompanied by the bubble growth and collapse. In hydrodynamics, this phenomenon is known as **cavitation.**

It should be noted that a pure liquid free of microscopic bubbles can withstand negative pressures (uniform extension without a discontinuity of the medium) of the order of 10^3 atm. However, real liquids containing microscopic gas bubbles can withstand negative pressures only of a few tens of atmospheres.

Cavitation in a fast liquid flow may lead to the formation of large cavities (caverns) in the regions of low pressure and their "self-liquidation" in the regions of elevated pressure. Cavitation affects the nature of the flow, causing a reduction in the lifting force and increasing the resistance of hydrofoils of sea liners. The materials of propellers, turbine blades, and high-speed pumps are subjected to erosion and destruction due to cavitation.

A quantitative analysis of cavitation can be based on the Bernoulli equation

$$p + \frac{\rho v^2}{2} = \text{const},$$

corresponding to the ideal incompressible liquid approximation. For steady-state flows, this equation is used for the introduction of dimensionless parameter **(cavitation number)**

$$\sigma = \frac{p - p_{cr}}{\rho v^2/2}, \qquad (6.12.1)$$

where p_{cr} is the critical pressure at which cavitation emerges. Normally, p_{cr} is taken equal to the saturated vapour pressure of the liquid. It is known that under this condition the liquid starts boiling.

Obviously, cavitation effects in a liquid with $p \approx p_{cr}$ are well pronounced. In this case, a single cavity **(cavern)** containing vapour and gases is formed behind a body. According to Eq. (6.12.1), such conditions in a flow past a body correspond to cavitation numbers σ close to zero. For small

σ, the size of cavern may considerably exceed the size of the body.

We can easily estimate the drag force F_{cav} acting on the formed cavitation cavern with a cross-sectional area S_{cav}. Using the general relation

$$F = C\rho v^2 S$$

derived by us above, we can write the following expression for the **cavitation drag coefficient:**

$$C_{cav} = \frac{F_{cav}}{\rho v^2 S_{cav}}. \qquad (6.12.2)$$

Considering that $[p - p_{cr}] \sim F_{cav}/S_{cav}$ in (6.12.1), we can write (6.12.2) in the form

$$C_{cav} = k\sigma,$$

where k depends on the type of the flow and assumes the value $k = \pi/4$ for two-dimensional flows, and $k = 0.9$-0.95 for axisymmetric flows.

Cavitation effects can be produced artificially by different methods, for example, by introducing a gas into the expansion region near a body. This is done to simulate cavitation effects in water tunnels not at a flow velocity exceeding 50 m/s, as is the case under natural conditions, but at a velocity of just a few metres per second. It is important in such a simulation that artificial and natural caverns should exist at the same cavitation number σ.[7]

6.13. Highly Rarefied Gases

The ideas about transport processes in gases like diffusion, heat conduction, and viscosity discussed in the previous sections are applicable when the mean free paths of molecules $l \ll L$. The quantity L can be regarded as the size of a vessel containing a gas or the size of a body moving in a gas.

Let us consider a gas when $l \gg L$. Such a gas is known as **highly rarefied.** We shall analyze the process of heat transfer in a highly rarefied gas. For this purpose, we can, for example, introduce into a vessel two plates having different

---

[7] For a complete similarity, the equality of the values of the Reynolds number Re and of some other criteria is also required.

temperatures T_1 and T_2 and separated by a distance L from each other.

It is well known that the mechanism of heat transfer in an ordinary gas is reduced to a random motion of energy carriers, viz. molecules. This is a "staggered path" through multiple intermolecular collisions. A different situation takes place in a highly rarefied gas: the intermolecular collisions are almost absent, and molecules are directly scattered by the plates. Having received an energy from the plate at a higher temperature and supplying a part of their energy to the plate at a lower temperature, molecules execute the energy transfer.

There is no sense to speak of a temperature gradient in the space between the plates in a highly rarefied gas. However, in analogy with the relation $j_E = -\lambda \mathrm{d}T/\mathrm{d}x$, we introduce

$$j_E = -\lambda \frac{T_2 - T_1}{L}. \tag{6.13.1}$$

How can we estimate λ for a highly rarefied gas? Proceeding from expression (6.5.11) $\lambda \sim nc_p vl$ in which we make the transition

$$l \to L, \tag{6.13.2}$$

we obtain

$$\lambda \sim nc_p vL. \tag{6.13.3}$$

It follows hence that for a highly rarefied gas, λ is not a characteristic of the gas itself since it is determined by the geometry of the vessel (through L).

Another argument in favour of what has been said above is that the thermal conductivity λ in ordinary gases does not depend on the number density n of gas particles. Indeed,

$$\lambda \sim nc_p vl \sim nc_p v \frac{1}{n\sigma} \sim c_p \frac{v}{\sigma}.$$

Here, we used the relation $l \sim 1/(n\sigma)$, where σ is the effective cross section of molecular interactions. In highly rarefied gases, the above dependence l (n) is not observed since relation (6.13.2), and hence (6.13.3), is valid. According to the latter relation, the change in the number density n of particles leads to a change in λ.

Practically, the situation described above can be as follows. We shall measure the thermal conductivity of the gas contained between two walls having different temperatures. Let us start measurements at the normal atmospheric pressure. The gradual pumping out of the gas from the space between the walls does not affect its thermal conductivity. However, as the gas becomes rarefied, we attain the moment when the mean free paths of the molecules become equal to the distance between the walls. With a further evacuation, the thermal conductivity rapidly drops.

This effect is used in heat-insulating **Dewar flasks.** A conventional thermos for storing hot or cold drinks, as well as the vessels for storing liquefied gases, have double walls with a high vacuum between them.

The internal friction in highly rarefied gases undergoes similar changes. Let us consider two solid surfaces moving relative to each other with a velocity u. Suppose that a highly rarefied gas is in the space between the surfaces. Then instead of Eq. (6.4.1) we can write

$$\Pi = -\eta \frac{u}{L}, \tag{6.13.4}$$

where Π is the friction acting on a unit surface of the plate, L is the separation between the plates, and η is the dynamic viscosity.

The estimate of η is based on (6.5.12), where l is replaced by L, i.e.

$$\eta \sim nmvL. \tag{6.13.5}$$

The drag force acting on a body moving in a highly rarefied atmosphere is given by

$$F \sim \Pi S \sim (nmv)uS, \tag{6.13.6}$$

where S is the area of the surface of the body, u is its velocity, and nmv is the density of molecular momentum imparted to the body.

Interesting effects can be observed in two communicating vessels containing highly rarefied gases at different temperatures. In this case, the mechanical equilibrium for ordinary gases will set in at the same pressure in the two vessels (only in this case will the gases not flow from one vessel to the

other). But how will the equilibrium set in in the case of highly rarefied gases?

Since the diameter of the connecting tube is small in comparison with the mean free path of molecules, the molecules will leave a vessel independently of one another. The molecules leaving the tube move with the same velocity as they entered it. In such a "molecular beam", the number of molecules leaving a vessel per second is Snv, where S is the cross-sectional area of the tube. Since $n = p/(kT)$ and $v \sim \sqrt{kT/m}$, we have

$$Snv \sim \frac{pS}{\sqrt{mkT}}. \tag{6.13.7}$$

The mechanical equilibrium in highly rarefied gases sets in when the numbers of molecules leaving the vessels per second equalize. In other words, it is observed when the total mass transfer in the two directions through the connecting tube is zero. As a result, we obtain

$$\frac{p_1}{\sqrt{T_1}} = \frac{p_2}{\sqrt{T_2}}. \tag{6.13.8}$$

It should be noted that for an ordinary gas, we have the relation

$$\rho_1 T_1 = \rho_2 T_2.$$

Formula (6.13.7) forms the basis of a method of isotope separation. An element is transformed into the gaseous state, the gas being a natural mixture of isotopes of the given element. An "atomic beam" emerges through a small hole in the vessel containing the rarefied gas mixture. According to (6.13.7), the beam is enriched in light isotopes.

6.14. Macroscopic Quantum Effects in a Liquid

Quantum effects start to be manifested in the properties of a liquid at temperatures close to absolute zero. Actually, we speak of a single liquid, viz. helium, or, to be more precise, its liquid isotope ^{4}He.

The condition for the existence of the liquid state can be written as follows:

$$\frac{\langle p^2 \rangle}{m} \sim \langle U \rangle, \tag{6.14.1}$$

where p is the momentum of a particle. This expression shows that at sufficiently low temperatures (the classical momentum p of a particle is proportional to the temperature) it is obviously violated. If, however, the mean energy $\langle U \rangle$ of interaction between the atoms of a substance is quite low, low temperatures can be attained such that (6.14.1) is replaced by

$$\frac{\hbar^2}{mr^2} \sim \langle U \rangle. \tag{6.14.2}$$

Here, we have used the quantum-mechanical uncertainty relations (4.3.2). In this formula, r is the interatomic distance which becomes of the order of the length of the probability wave associated with the thermal motion of the atoms. For liquid helium, the binding energy per atom $\langle U \rangle = -\text{const}/r^6$ (see the end of Sec. 4.10) is equal to 2.5×10^{-3} eV. The numerical estimate of the expression $\hbar^2/m_{\text{He}}r^2$ for helium atoms at $r \sim 10^{-10}$ m gives the same result. Thus, condition (6.14.2) is valid for liquid helium up to the absolute zero since temperature does not appear in Eq. (6.14.2).

The fact that liquid helium does not freeze at absolute zero is a quantum phenomenon. Moreover, ^{4}He nuclei have integral spin (unlike ^{3}He nuclei), which means that ^{4}He atoms are Bose particles (see Sec. 4.6.2). At $T = 0$ K, a considerable fraction of such particles can be in the ground (lowest) quantum state. In other words, any macroscopic motion of such a liquid that does not involve its transition from the ground state is not accompanied by a change in the energy and momentum of the liquid as a whole, i.e. the liquid has no viscosity **(superfluidity)**.

How is hydrodynamics of a quantum liquid constructed? Since, as was mentioned above, the quantum liquid is a system in which a macroscopic fraction of particles is in the ground quantum state, it is natural to assume that a superfluid liquid as if consists of two mutually penetrating liquids: the normal component that behaves like a classical (ordinary viscous) liquid and the superfluid component which

is an aggregate of a macroscopic number of particles moving coherently in a single quantum state. Here, we must assume that "there is no friction" between these two parts of the mass of the liquid, i.e. there is no momentum transfer between the components.

In order to describe the motion of a quantum liquid, eight variables are required: the density ρ_n of the normal component, the pressure p, the velocity $\mathbf{v}_n$ of the normal component (see Sec. 6.2), and in addition, the velocity $\mathbf{v}_s$ of the superfluid component (here, we assume that the density ρ_s of the superfluid component is determined from the equation $\rho_s = \rho - \rho_n$, where ρ is the total density of the liquid). Since the equation of motion for a quantum liquid contains the two velocities, it is described by a **two-velocity hydrodynamics**. For the same reason, it is more appropriate to speak of two types of motion in the quantum liquid, each of which is associated with its own effective mass. Naturally, the effective masses of densities ρ_n and ρ_s are functions of temperature ($\rho_n = 0$ and $\rho_s = \rho$ at $T = 0$ K).

Let us consider some properties of the flows of ^{4}He quantum liquid. No viscosity is observed in experiments with a liquid helium flow through capillaries and slits (with a width of 10^{-6} m). However, in experiments on measuring the viscosity of a liquid with the help of torsional vibrations of a disc immersed in the liquid, finite values of friction are obtained. Clearly, the superfluid motion in a liquid is observed in experiments with capillarics, while the damping of disc vibrations confirms the presence of the normal motion.

The existence of two types of motion in liquid helium[8] is demonstrated most effectively in an experiment on rotation of a cylindrical vessel filled with ^{4}He. The walls of the rotating cylinder entrain only the normal component of the liquid, while the superfluid component remains at rest.

Another peculiarity of the superfluid flow in a quantum liquid is its inability to heat transfer. As a matter of fact, the superfluid component of the effective mass associated with the corresponding motion is completely unexcited as if it were at absolute zero. Elementary thermal excitations

[8] We speak of liquid helium that has undergone the second-order phase transition. Helium condenses at $T = 4.2$ K. The liquid-liquid transition mentioned above takes place at the critical temperature $T_{cr} \simeq 2.2$ K of transition.

form the normal component of the corresponding effective mass, and the normal macroscopic motion has the mechanism of heat conduction in a quantum liquid. Hence it becomes clear why the thermal conductivity of liquid helium has the record values.

This peculiarity of liquid helium is visually manifested in the **thermomechanical effect** consisting in the heating of a part of liquid remaining in a vessel from which helium flows out through a capillary.

Thus, the superfluid flow of ^{4}He quantum liquid has neither viscosity nor thermal conductivity, i.e. we have all the properties of an ideal liquid whose flows must be vortex-free. A body placed in an ideal liquid flow does not experience a drag.

As certain critical velocities are attained in liquid helium, quantized macroscopic vortices emerge in it. As a matter of fact, the angular momentum of any quantum-mechanical system (in the case under consideration, the whole liquid) can change by an integral value. Therefore

$$m_{\mathrm{He}} v r = n\hbar, \tag{6.14.3}$$

where $n = 1, 2, 3, \ldots$, r is the distance between an atom and the axis of a vortex, and v is the linear velocity of the atom in the vortex. A quantized vortex may embrace the entire volume of the liquid.

Equations of motion for a rotating liquid helium contain Planck's constant $\hbar$. Normally, this constant appears in the relations describing microscopic objects. Thus, the motion of a superfluid liquid is a demonstration of quantum effects on the macroscopic scale.

6.15. Generalizations of Hydrodynamics

One of the generalizations of hydrodynamics, viz. two-velocity hydrodynamics of a quantum liquid, was considered earlier.

Recently, a new branch has appeared in nuclear physics, i.e. **relativistic nuclear physics**. It deals with the properties of nuclear matter at extremely high energies. The very concept of nuclear matter has a meaning only for heavy nuclei containing hundreds of nucleons. Highly strong excitations of such nuclei may appear as a result of their collisions.

Theoretical estimates show that the "velocity of sound" for

nuclear matter, determined in terms of its "compressibility", is (0.1-0.2)c. If the relative velocity of colliding heavy nuclei exceeds the velocity of propagation of sound in nuclear matter, the compression of a compound superheavy nucleus is similar to a shock wave in a macroscopic medium. Experiments proved that hydrodynamic effects do exist during supersonic collisions of heavy nuclei. These effects can be described by relativistic generalization of hydrodynamics.

The above experiments aimed at the quest for new forms of nuclear matter that may appear under extremal conditions of very high temperatures and high degrees of compression. We are speaking about phase transitions in nuclear matter, accompanied by superdense states of the nuclear matter. Extremal conditions were created experimentally during the collisions of heavy ions accelerated to relativistic energies. As a result, the states of nuclear matter with a temperature of the order of 10^{12} K have been obtained. However, no methods have been worked out so far for determining the density of nuclear matter during a collision, although a compression ratio of several times can be obtained for an energy of incident nucleus of 10^8 eV per nucleon. However, two quantities, viz. temperature and density, are required to write the equation of state for nuclear matter.

The equation of state for nuclear matter is essential in cosmology, in investigation of a gravitational collapse,[9] in the study of neutron stars, and for the development of nuclear physics proper (see Chap. 9).

In astrophysics, such large-scale phenomena as bursts of supernovas are known. They can also be described only with the help of concepts of relativistic hydrodynamics and the corresponding theory of shock waves. In Sec. 6.10.4, we considered the problem on a high-intensity explosion in the atmosphere. It would be quite correct to apply its solution to the description of the bursts of supernovas. Such a description does not reveal the mechanism of the explosion proper, but rather explains phenomena accompanying it.

An analysis of the behaviour of a highly ionized gas (plasma) in magnetic fields has led to the development of **magnetohydrodynamics** (see Sec. 8.6).

[9] A collapse is an uncontrollable falling of a substance of bulk astrophysical objects to their centre after the internal sources of nuclear energy have been exhausted.

Chapter 7

ELECTROMAGNETIC FIELDS IN MEDIA. ELECTRICAL, MAGNETIC, AND OPTICAL PROPERTIES OF SUBSTANCES [1]

7.1. Superconductivity

A remarkable property is observed for a number of metals and alloys at temperatures of several kelvins. Their **resistance** drops to zero (resistance is reciprocal to electrical conductivity). This phenomenon, called **superconductivity**, could not be explained almost for fifty years (it was discovered in 1911). The development of quantum mechanics in the twenties could not bring a light to this phenomenon, although it laid the principal foundations for its explanation. The mechanism of this phenomenon became clear to theoreticians only after profound experiments had been carried out to investigate this phenomenon.

The most important experiment in this field includes the discovery of the **isotope effect** in superconductors. *If the isotopic composition of a superconducting material changes, the critical temperature T_{cr} of transition from the normal state of a metal to the superconducting state also changes.* The velocity of sound in a superconductor depends on the atomic mass of its material. Therefore, it is natural to assume that superconductivity is somehow related to the interaction between electrons and phonons, i.e. quanta of sound (see Sec. 5.10.5). Indeed, under the action of an electron, the lattice is deformed. This deformation, in turn, affects another electron. As a result, these two electrons interact through a phonon (in addition to the ordinary Coulomb interaction).

The Coulomb interaction of electrons leads to their mutual repulsion, while the interaction of electrons through lattice vibrations excited by them leads to their attraction. At very low temperatures, the effective attraction of electrons in some metals dominates over their repulsion, and the pairs of "bound electrons" are formed. Naturally, the energy of pairs is lower than the total energy of free electrons. As a result, a "gap" is formed near the Fermi energy ε_F (see

[1] CGS units are used in this chapter.

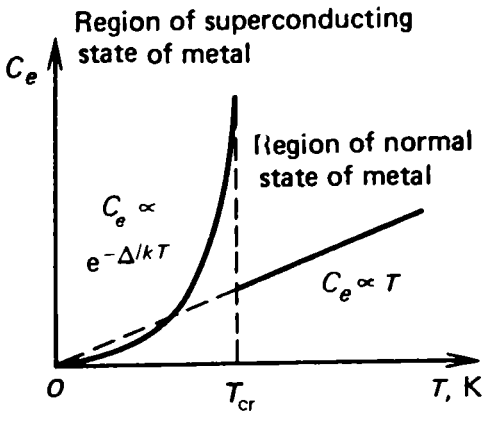


FIG. 7.1

Sec. 5.10.6) in the energy spectrum of electrons between the ground and excited states. The fact that an "energy gap" is formed in superconductors follows from the analysis of the behaviour of the electron heat capacity C_e (Fig. 7.1). By the way, the results of this analysis also indicate that the superconducting transition is a second-order phase transition (see the end of Sec. 5.11).

Electron pairing takes place in the states with total zero momentum and zero spin. Paired electron states are Bose formations unlike individual electrons which have a half-integral spin and obey Fermi statistics. Since paired electrons behave like Bose particles, they can be unlimitedly accumulated in the ground quantum state which behaves macroscopically. A current once excited in this state without destroying the state itself will flow eternally (due to the presence of the energy gap) without any external source of energy. This is just superconductivity. This phenomenon can be regarded as the superfluidity of charged electron liquid which in this case is a Bose liquid (see Sec. 6.14).

It should be noted in conclusion that at temperatures T_{cr} high enough to break electron pairs, a metal goes over from the superconducting to the normal state (see Fig. 7.1). The quantity kT_{cr} is of the order of the width Δ of the energy gap. In most cases, the temperature does not exceed 20 K (for niobium-tin alloy).

7.2. Electrical Conductivity of Metals

The state of a free electron is determined by specifying only its momentum **p**, the energy being a function of momentum ($E = p^2/(2m)$). According to the uncertainty relations (4·3.2), by specifying the conduction electron momentum, we make its coordinates indeterminate, i.e. all points of location of the electron in a lattice become equiprobable. In

other words, the lattice is transparent for conduction electrons, and the resistance of an ideal lattice should be zero (this situation should not be confused with superconductivity).

Since for an electron in a lattice the space has the symmetry of the lattice, it must be described not by the actual momentum but by the quasi-momentum (cf. similar considerations for phonons in Sec. 5.10.5). This means that the probability waves describing the conduction electron distribution in the periodic field of the lattice must possess the symmetry of the lattice itself, i.e.

$$\psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{a}),$$

where **a** is the lattice constant.

There are several factors that make impossible the equality to zero of the electrical resistance of metals which are not in a superconducting state even at $T = 0$ K. Firstly, crystal lattices cannot be strictly ideal. They always contain some kinds of defects like impurities or vacancies. This leads to the "scattering" of the probability waves by lattice inhomogeneities. Secondly, the existence of thermal motion at $T \neq 0$ K leads to the violation of the strict periodicity of the lattice, and hence to the scattering of electron waves by lattice vibrations. These factors make the resistivity ρ a complex function of temperature T. A typical $\rho(T)$ dependence is shown in Fig. 7.2.

Obviously, defects and thermal irregularities of the lattice exist in the superconducting state of a metal as well, but

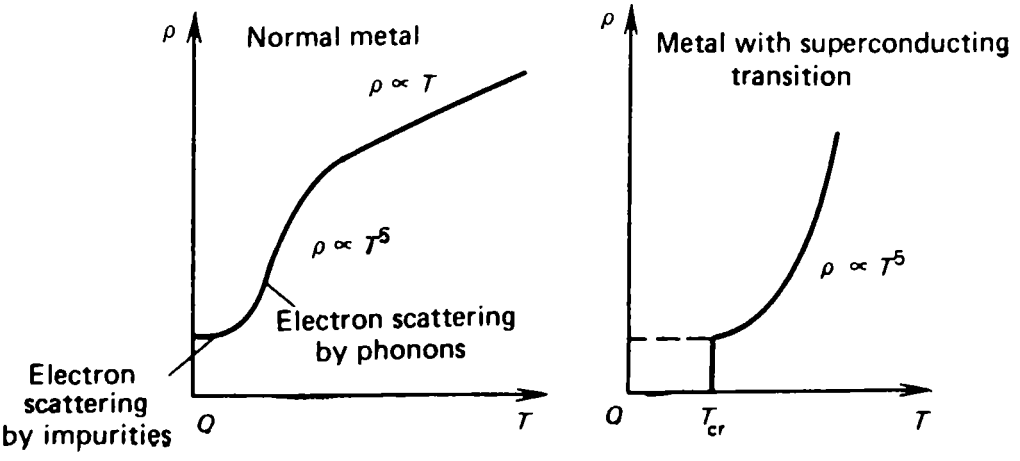


FIG. 7.2

they are insignificant. As a matter of fact, the ground state which is occupied by practically all paired electrons exhibits a sort of "stiffness" relative to such perturbations as structural and thermal irregularities of the lattice. Indeed, the probability of a transition of a Bose particle to the lowest quantum state ("condensate") is proportional to the number of particles in this state, which is huge. All the particles of the "Bose condensate" are in the same state, and if the "condensate" moves as a whole forming a current, it is rather difficult to take a Bose particle out of this state. On the contrary, although in a normal metal, electrons are collectivized in the conduction band, each of them is in an individual quantum state (Fermi particles). Any perturbation causes the scattering of such an independent electron, and it comes out of play.

Let us now analyze the electrical conductivity of normal metals quantitatively. For the sake of mathematical simplicity, we shall carry out the analysis by using the semiclassical model of electron gas. The expressions obtained in this case differ from their quantum-mechanical counterparts in the values of numerical factors but coincide with them in the order of magnitude.

In the absence of an external electric field **E**, the electron gas in an ionic lattice is in thermal equilibrium with it. The application of an electric field to a crystal creates a nonequilibrium electron distribution in the lattice. The establishment of equilibrium in the system is characterized by the emergence of an electron current. The current density $\mathbf{j}_e$ is connected with the electric potential gradient $\overrightarrow{d\varphi}/dr$ through the relation

$$\mathbf{j}_e = -\sigma \frac{\overrightarrow{d\varphi}}{dr}, \qquad (7.2.1)$$

where σ is the electrical conductivity. Relation (7.2.1) is similar to Eqs. (6.4.1)-(6.4.3). This is not surprising since viscosity, heat conduction, and diffusion are also the mechanisms of transition to an equilibrium state when the corresponding gradients of velocity, temperature, or number density emerge in a system.

Expression (7.2.1) can be written in the form

$$\mathbf{j}_e = \sigma \mathbf{E}, \qquad (7.2.2)$$

where **E** is the electric field strength. While deriving this expression, we have used formula (3.4.14).

Relation (7.2.2), like (7.2.1), is a macroscopic and essentially approximate relation if we apply it not only to metals but to arbitrary conductors of current. In metals, relation (7.2.2) is satisfied virtually for all values of electric field strength. In gaseous conductors (plasma), liquid conductors (electrolytes), and semiconductor crystals, relation (7.2.1) holds only for small potential gradients.

Let us continue the analysis of normal metals. We shall derive the expression for the current density by using a microscopic approach. We proceed from the obvious relation

$$\mathbf{j}_e = en\mathbf{v}_{dr} \tag{7.2.3}$$

where e is the electron charge, n is the electron number density, and $\mathbf{v}_{dr}$ is the **drift velocity** (the velocity of electron displacement in an external electric field). This velocity is superimposed on the thermal velocity of electrons and is extremely small in comparison with the latter. In turn, the drift velocity $\mathbf{v}_{dr}$ of electrons in an electric field **E** is determined from the equation of motion for charges:

$$m\frac{\mathbf{v}_{dr}}{\tau} \sim e\mathbf{E}, \tag{7.2.4}$$

where m is the electron mass, and τ is the mean free time of the electron, which is of the order of relaxation time (see Secs. 6.1 and 6.5.1). Substituting the value of $\mathbf{v}_{dr}$ from (7.2.4) into (7.2.3), we obtain

$$\mathbf{j}_e \sim \frac{e^2 n\tau}{m}\mathbf{E}. \tag{7.2.5}$$

Let us now compare the microscopic (7.2.5) and the macroscopic (7.2.2) expression for current density. It leads to the following expression for the electrical conductivity:

$$\sigma \sim \frac{e^2 n}{m}\tau. \tag{7.2.6}$$

Thus, we managed to express the kinetic coefficient σ, responsible for the charge transfer, in terms of the parameters of elementary charge carriers.

The electrical conductivity σ can also be expressed in terms of the mean free path l by using a relation similar to (6.5.1*), i.e.

$$\tau v \sim l, \tag{7.2.7}$$

where v is the thermal velocity of an electron, which is of the order of the Fermi velocity v_F (see Eq. (5.10.22)):

$$v_F = \frac{p_F}{m} = \frac{\hbar n^{1/3}}{m}. \tag{7.2.8}$$

As a result, Eq. (7.2.6) can be written as

$$\sigma \sim \frac{e^2 n}{m v_F} l. \tag{7.2.9}$$

A conduction electron in a lattice participates in various scattering processes: (i) electron-phonon scattering, (ii) electron scattering by defects, and so on. Each type of fundamental interaction is characterized by the effective scattering cross section (see Sec. 6.5.1). The total scattering cross section

$$\sigma_\Sigma = \sigma_{eff} + \sigma_{e.d} + \ldots,. \tag{7.2.10}$$

where σ_{eff} is the effective cross section of electron scattering by collective quantum vibrations of the lattice, i.e. electron-phonon scattering, and $\sigma_{e.d}$ is the cross section of electron scattering by lattice defects. Naturally, the cross sections of electron scattering on inhomogeneities of different types (empty lattice points, i.e. vacancies, or on impurities of foreign atoms) are slightly different. In Eq. (7.2.10), the contribution made by a certain type of electron scattering depends on the temperature region. For temperature close to absolute zero, when the thermal motion of the lattice is insignificant, the main contribution is made by the electron scattering by defects (the residual resistance of a metal at $T = 0$ K). An insignificant increase in temperature is accompanied by a rather rapid growth of the thermal excitation of the lattice, and the electron scattering by phonons becomes predominant. The described pattern is graphically represented in Fig. 7.2 which shows the temperature dependence of $1/\sigma$ (reciprocal of electrical conductivity).

In order to make everything clear, we transform Eq. (7.2.10) as follows. We proceed from the relation (6.5.1)

between the effective scattering cross section and the corresponding mean free path l. Then Eq. (7.2.10) can be written in the form

$$\frac{1}{l} = \frac{1}{l_{\text{eff}}} + \frac{1}{l_{\text{e.d}}}. \tag{7.2.11}$$

As a result, in expression (7.2.9) for electrical conductivity, the corresponding expression from (7.2.11) will appear instead of l. In Eq. (7.2.9), e and m are constants. According to Eq. (7.2.8), v_{F} practically does not depend on temperature, while according to (5.10.25), n is a weakly increasing linear function of T. Therefore, the dependence $l(T)$ is decisive for the temperature variation of electrical conductivity.

It would be interesting to compare the electron thermal conductivity of metals with their electrical conductivity. For this, we shall write the ratio of the thermal conductivity λ to the electrical conductivity σ. We shall use the general kinetic equation (6.5.11) for λ, and Eq. (7.2.9) for σ. This gives

$$\frac{\lambda}{\sigma} \sim \frac{C_e v_{\text{F}} l}{\left(\frac{e^2 n}{m v_{\text{F}}} l\right)} = \frac{C_e m v_{\text{F}}^2}{e^2 n}. \tag{7.2.12}$$

Let us consider the model of a purely classical electron gas. This model is applicable to metals only for temperatures exceeding the electron degeneracy temperature (see Eq. (5.10.19)). In the gaseous plasma, this model is quite feasible. In the framework of the classical analysis (see Secs. 5.8.3 and 5.8.4), $mv^2 \sim kT$ and $C \sim nk$. Then relation (7.2.12) can be written as

$$\frac{\lambda}{\sigma T} \sim \left(\frac{k^2}{e^2}\right). \tag{7.2.13}$$

The right-hand side of this relation is constant. Relation (7.2.13) is known as the **Wiedemann-Franz law**.

7.3. Direct Current

Macroscopic relation (7.2.1) is known as **Ohm's law**. In order to write it in the conventional form, we apply (7.2.1) to a straight segment of length L of a metal conductor of cylindrical shape (Fig. 7.3). Let current I of density $\mathbf{j}$ be

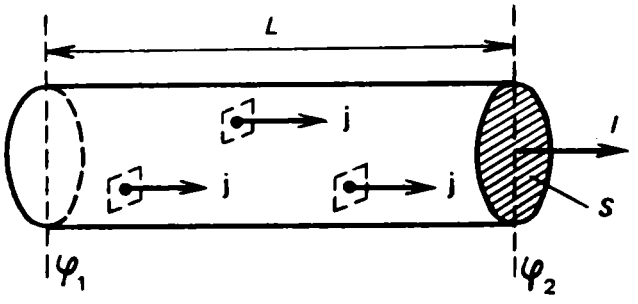


FIG. 7.3

uniformly distributed over the geometrical cross section S of the conductor and directed along the symmetry axis of the cylinder. Then we can write the following expression for the potential gradient in (7.2.1):

$$\left|\frac{\overrightarrow{\mathrm{d}\varphi}}{\mathrm{d}r}\right| = \frac{\varphi_2 - \varphi_1}{L}. \tag{7.3.1}$$

Since the current density $\mathbf{j}_e$ has the dimensions $\mathrm{C/(m^2 \cdot s)}$, the total charge flux through the cross section S is

$$I = j_e S = \frac{\mathrm{d}\sum e}{\mathrm{d}t}. \tag{7.3.2}$$

The quantity I is called the **current.**

Let us write Eq. (7.2.1) in the form

$$j_e = -\sigma\left|\frac{\varphi_2 - \varphi_1}{L}\right|. \tag{7.3.3}$$

Multiplying both sides of this equation by S, considering that the potential φ drops along the direction of the field, and also using (7.3.2) and the well-known relation

$$\sigma = \frac{1}{\rho}, \tag{7.3.4}$$

we obtain

$$I = \frac{\varphi_2 - \varphi_1}{\rho L/S} = \frac{\Delta\varphi}{R}, \tag{7.3.5}$$

where

$$R = \frac{\rho L}{S}, \tag{7.3.6}$$

ρ being the resistivity of the conductor material, and R the total resistance of the conductor. Formula (7.3.5) is the conventional form of Ohm's law, which is the basic relation in an analysis of d.c. circuits.

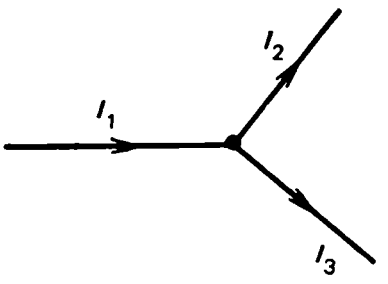


FIG. 7.4

In the general case of branched d.c. circuits, it is convenient to carry out their analysis with the help of **Kirchhoff's laws**. The first law is an expression of the charge conservation law for electric circuits. Let us consider a junction of an arbitrary electric circuit (Fig. 7.4). From the charge conservation law, we immediately obtain

$$I_1 = I_2 + I_3,$$

or

$$I_1 - I_2 - I_3 = 0.$$

In the general case, the last equality can be written as

$$\sum_i I_i = 0. \tag{7.3.7}$$

The algebraic sum of currents in any junction of a circuit must be zero.

The second law is just the condition of the steady-state process of passage of a direct current. It states that *the sum of all the potential differences along a closed circuit is equal to the potential difference across the source.* Let us consider Fig. 7.5. The above condition can be quantitatively written as

$$\sum_{\substack{\text{along}\\ \text{circuit}}} (\varphi_i - \varphi_k) = \Delta\varphi_{AB}, \tag{7.3.8}$$

where $\Delta\varphi_{AB} = \varphi_A - \varphi_B$, or

$$\sum_{\substack{\text{along}\\ \text{circuit}}} (\varphi_l - \varphi_k) = \frac{e\Delta\varphi_A}{e} \equiv \mathscr{E}. \tag{7.3.9}$$

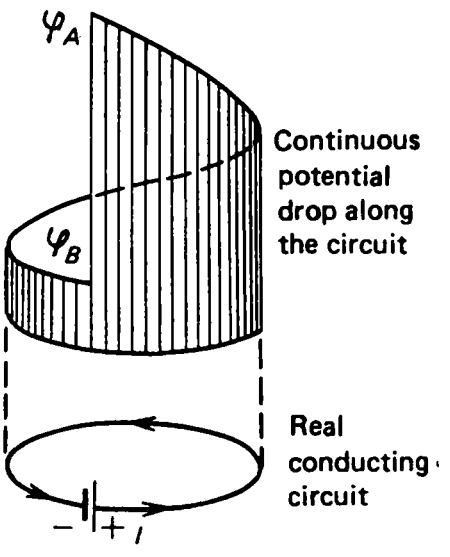


FIG. 7.5

Using Ohm's law in the form (7.3.5), we can write Eq. (7.3.9) as

$$I \sum_i R_i + Ir = \mathscr{E}, \qquad (7.3.10)$$

where R_i is the external resistance of individual subsections of the circuit, r is the internal resistance of the source, and $\mathscr{E}$ is the emf of the source. The physical meaning of the latter quantity follows from relations (7.3.9) and (3.4.13). In the case of simple circuits, relations (7.3.9) and (7.3.10) are applied to an entire closed circuit. For branched circuits, these relations are written in the form of a system of equations for all possible subcircuits of the circuit. The subcircuits may contain or not contain current sources.

While calculating branched circuits, *the number of equations of the type* (7.3.7) *for junctions and the number of equations of the type* (7.3.9) *for closed subcircuits must be equal to the number of the unknown parameters of the circuit.*

Let us consider the last remark in greater detail. The current in the circuit caused by the field of the source carries the energy flux

$$J = \frac{\mathrm{d}\sum e}{\mathrm{d}t} \Delta\varphi = I\,\Delta\varphi = I^2R. \qquad (7.3.11)$$

In the last link of this chain of equalities, we have used Eq. (7.3.5). Expression (7.3.11) is the energy transferred through the whole cross section of a conductor per second, i.e. J is the power of the current.

The passage of a current through a conductor is accompanied by thermal effects. Therefore, the expression I^2R is sometimes referred to as **Joule losses.**

7.4. Dielectric Conductance

7.4.1. Electrons and Holes. Exciton States. In insulating crystals, unlike in metals, electrons do not populate the conduction band at $T = 0$ K (see Fig. 5.26). The conduction band in these materials is separated from the filled band by an energy gap Δ. Hence it is clear that at absolute zero all dielectrics are insulators (their resistivity $\rho = \infty$). However, at $T \neq 0$ K, the electrical conductivity of insulating crystals strongly depends on the energy-gap width. Indeed, the number of charge carriers crossing the forbidden band due to thermal excitation and emerging in the conduction band is proportional to $\exp -[\Delta/(kT)]$. Therefore, crystals with a wide forbidden band belong to **insulators**, while those with a narrow forbidden band belong to **semiconductors.**

Diamond is an example of a good insulator. Its energy-gap width $\Delta = 6$-7 eV. Typical semiconductors include germanium ($\Delta = 0.72$ eV) and silicon ($\Delta \simeq 1.1$ eV).

Since 1 eV/k = 11 600 K, a thermal transition of electrons through the gap in the energy spectrum is ruled out in insulators. However, such a transition can be realized by exposing an insulator to electromagnetic radiation or high-energy particles. For example, the radiant energy required to transfer an electron from the filled band to the conduction band for insulators with an energy-gap width of the order of 10 eV must be not less than 10 eV. Since, according to Eq. (4.7.3), the energy of a photon is $E = \hbar\omega = 2\pi\hbar c/\lambda$, for the wavelength of the required electromagnetic radiation we obtain $\lambda \simeq 100$ nm. This wavelength corresponds to ultraviolet radiation.

An external electric field applied to an insulator is capable of accelerating photoelectrons which emerge in the conduction band, thus producing a current. Visible light normally cannot induce photoconduction in insulators. As was shown above, ultraviolet radiation is required for this purpose.

Metals are known to exhibit only **electron conduction**, while in semiconductors and insulators, it is supplemented by **hole conduction**. As a matter of fact, when an electron goes over from the filled band to the conduction band, a vacancy ("hole") is formed in the filled band. As a result,

the electrons of this lower band can be redistributed, forming a hole current in an external field. Therefore, current emerges in the two bands simultaneously.

As regards the signs of the charge carriers in the two bands, they must be opposite. This follows from the charge conservation law and the fact that a crystal as a whole remains electrically neutral (its total charge is zero).

In view of the Coulomb interaction between a "positively charged hole" and a negative electron, they may form a bound state. Such a state is known as the **exciton**. The exciton resembles the hydrogen atom and has a similar energy spectrum (see Fig. 4.10). The discrete negative levels of this spectrum, which become denser near the zero value, correspond to excitons with different degrees of excitation. Positive quasi-continuous energy values in the exciton spectrum correspond to a free electron and a hole.

It should be borne in mind that the exciton state is of collective nature, characterizing a crystal as a whole. This follows from the collective nature of the electron and hole states in the crystal. The exciton state propagates in the form of an exciton wave over the crystal. For the sake of clarity, this situation is usually interpreted as an energy transfer from one excited atom to another. When the excitation energy is sufficient to ionize an atom, a free electron and a positively charged ion are formed.

The concept of the exciton state is very fruitful for the investigation, for example, of biological structures on molecular level.

7.4.2. Semiconductors. In crystals with a narrow energy gap (semiconductors), the conductivity sharply increases with temperature. For this reason, they can be used as **thermistors.**

Semiconducting crystals like germanium or silicon have loose lattices which are ready to admit alien atoms, viz. impurities. By introducing impurities, it is possible to increase the conductivity of semiconductors by a few orders of magnitude.

Semiconductors exhibit two types of conduction: electron (*n*-type) and hole (*p*-type). Accordingly, impurity atoms may supply extra electrons or holes. Impurities that readily

give away their electrons are known as **donors**, while those capturing electrons are referred to as **acceptors.**

Let us consider a germanium crystal. It has a diamond-type lattice. Germanium atoms are tetravalent. Let us introduce pentavalent arsenic atoms into the germanium lattice. Then an arsenic atom occupying a vacant site of the lattice will spend four of its five valence electrons to form the bonds in the lattice, while the fifth electron remains "out of service". The binding energy of this unemployed electron is less than 0.1 eV. Thermal fluctuations (even at room temperature) are quite sufficient for the extra electron of the arsenic atom to go over to the free state and become a charge carrier. As a result, the arsenic atom becomes a donor impurity.

We can also introduce a trivalent impurity atom into the germanium lattice, say, an aluminium atom. The aluminium atom can borrow a missing electron from germanium itself. As a result, a hole is formed in the germanium atom, which starts to roam over the crystal as a positive charge carrier (as a matter of fact, an atom with a "hole" is a positive ion). Thus, the aluminium impurity becomes a supplier of holes, i.e. an acceptor.

In a pure (free of impurity) semiconductor, the numbers of positive and negative charge carriers are equal. In a semiconductor with a donor impurity, there is an excess of negatively charged carriers, and we have an *n*-type semiconductor. A semiconductor with an acceptor impurity is characterized by an excess of positively charged carriers and is referred to as a *p*-type semiconductor.

Interesting phenomena occur at the boundary between two regions of a semiconductor with different types of conduction. In such semiconductors with a *p-n* junction, a potential difference emerges across the junction (whose width is about 10^{-6} m). This can easily be explained: the number densities of the carriers of one sign are different on both sides of the *p-n* junction, which must cause a diffusion charge flow. This flow violates the electrical neutrality of the regions, and an electric field emerging over the width of the *p-n* junction prevents from a further diffusion of charges. Naturally, the diffusion flow is absent in the state of thermodynamic equilibrium, while the potential difference across the *p-n* junction differs from zero.

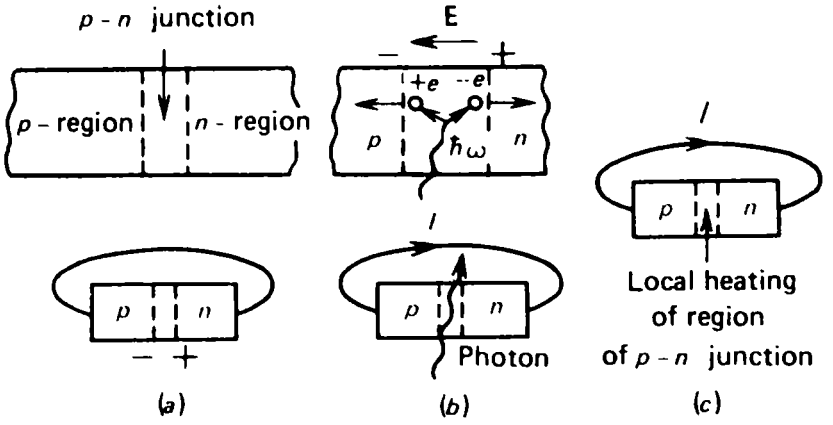


FIG. 7.6

If we short the ends of a semiconductor sample with a *p-n* junction (Fig. 7.6*a*), there should be no current in the circuit since the system as a whole is in thermodynamic equilibrium. If, however, we expose the *p-n* junction to light (Fig. 7.6*b*), electron-hole pairs will be formed in it, which will be immediately separated by the field (a hole moves to the *p*-region, and an electron, to the *n*-region). A current will appear in the system which is no longer in equilibrium. The radiant energy of sunlight is transformed into electrical energy. This principle forms the basis of operation of solar cells which are widely used on spacecrafts.

Figure 7.6*c* illustrates the operation principle of a thermopile. The only difference from solar cells consists in that nonequilibrium charge carriers in thermopiles are produced by a heat flow in them.

In any source of constant emf, charge carriers of both signs are separated at the expense of external energy (e.g. light, heat, and energy of chemical reactions). As a result, a potential difference is formed, which creates and maintains a current in the circuit.

7.5. Electric Fields in Matter

7.5.1. Field Fluctuations in a Substance. Let us mentally isolate in a metal a small volume containing a sufficiently large number of free electrons. As a result of random thermal motion of charges in the electron gas of the metal, the number of electrons in the isolated volume continuously fluctuates about the mean value. Naturally, these fluctua-

tions in the electron number density are responsible for the emergence of fluctuating electromagnetic fields. It should be borne in mind that the substance as a whole is obviously electrically neutral, and we speak of very small deviations from the electrical neutrality in the arbitrarily chosen element in the volume of the body. Such fluctuating spontaneously emerging fields are always present in a substance. An external electric field $\mathbf{E}_{ext}$ in a conductor is superimposed on the fluctuating internal field $\Delta\tilde{\mathbf{E}}$. Thus, the mean field in a substance is

$$\langle \mathbf{E} \rangle = \mathbf{E}_{ext} + \Delta\tilde{\mathbf{E}}. \tag{7.5.1}$$

Normally, the fluctuations of macroscopic quantities (see Sec. 5.2) present no interest in view of their smallness. However, sometimes situations arise when an observed effect is due to fluctuation phenomena. For example, the sensitivity threshold of high-precision electrical measuring instruments is determined by **noise currents**. As a matter of fact, weak random spontaneous currents emerge in a circuit owing to fluctuations even in the absence of an external emf.

Let us estimate the magnitude of such noise currents. The power associated with a pulse of fluctuation current can be written as follows:

$$I^2R \sim \frac{kT}{\tau}, \tag{7.5.2}$$

where τ is the relaxation time. Naturally, the mean square of the fluctuation current must be determined over the time $\Delta t \gg \tau$. Therefore, from Eq. (7.5.2) we obtain

$$\langle I^2 \rangle \sim \frac{kT}{R\,\Delta t}. \tag{7.5.3}$$

Current fluctuations in a circuit are the smaller, the higher the resistance and the lower the temperature.

If we take a metal conductor with a certain atomic composition, its electrical resistance R is proportional to the total number N of particles in the body:

$$\sqrt{\langle I^2 \rangle} \propto \frac{1}{\sqrt{N}}. \tag{7.5.4}$$

This result is in agreement with Eq. (5.2.3).

7.5.2. Electrostatic Fields in Metals. An external electrostatic field in a metal is compensated (the mechanism of the compensation will be described below). Therefore, in the bulk of the metal

$$\langle \mathbf{E} \rangle = 0 \tag{7.5.5}$$

since otherwise currents would be created, bringing about inevitable Joule losses. A steady-state current can be sustained by a permanent energy source which is absent in the case under consideration.

When a metal is introduced into an external electrostatic field, a certain movement (redistribution) of its free electrons takes place at the initial moment. These electrons are arranged on the surface of the metal body and screen the electrostatic field of external charges, thus ensuring the fulfilment of equality (7.5.5).

The external electrostatic field is normal to the surface of the metal since with any other configuration of the field, it would have a tangential component $\langle \mathbf{E}_\tau \rangle$ to the surface, and hence a surface current. But as was proved above, this is impossible. Therefore,

$$\langle \mathbf{E}_\tau \rangle = 0. \tag{7.5.6}$$

Let us consider a corollary of this equality. We proceed from relation (3.4.14) which expresses the electrostatic field strength in terms of the derivative of potential along a certain direction. Taking into account Eq. (7.5.6), we may conclude that the *potential* φ *along the surface of a metal remains constant.*

The component E_n of the electrostatic field strength, normal to the surface, is determined by the surface charge density σ. This relation follows from the fundamental field equation (3.3.3). The examples of calculation of the electrostatic field E_n on metal surfaces in the form of a sphere, line, plane surface, and parallel-plate capacitor were considered in Sec. 3.4.2. It should be borne in mind that at points which are not too close to the surface of a body, the mean field $\langle \mathbf{E} \rangle$ in vacuum actually coincides with the true field $\mathbf{E}$. The values of these two quantities differ only in the immediate vicinity of the body where the effects of irregular atomic fields (field fluctuations of the medium) are pronounced.

In Sec. 3.4.2, the field was determined from a given distribution of charges producing it on the surface of a metal. Here, we consider the inverse problem: a metal is in a given external electrostatic field, and we must find the corresponding static (constant) distribution of free charges in the metal.

The existence of a direct and an inverse problem indicates that the charges and potentials of a metal surface cannot simultaneously be specified arbitrarily (instead of the value of the field on the surface, we speak of a constant potential of the surface). In view of the linearity of the field equation (3.3.3), the dependence of the total charges Q_a on the potentials φ_a of the metal surfaces must also be linear, i.e. expressed through a relation of the form

$$Q_a = \sum_i C_{ai}\varphi_i, \tag{7.5.7}$$

where a is the number of a body, and the summation is carried out over all the bodies of the system, including the body under consideration. The constants C are **capacitances** which have the dimensions of length and depend on the shape and mutual arrangement of the metal bodies. In particular, if the system contains only one body, we have

$$Q = C\varphi, \tag{7.5.8}$$

and the order of magnitude of the capacitance C coincides with the linear dimensions of the body.

7.5.3. Electrostatic Fields in Insulators. Polarization of a Substance. Since an insulator contains no free electrons, the arguments used for describing metals are inapplicable here. Indeed, in the state of equilibrium, the mean field $\langle \mathbf{E} \rangle$ in the insulator may differ from zero since no current can appear.

Since insulators include molecular, ionic, and covalent crystals (see Secs. 5.10.2 and 5.10.4), it is possible that from the electrical point of view, their atomic or molecular structures are dipoles (see Sec. 3.4.4). In some cases, atomic dipole moments may orient (polarize) one another, and the regions of **spontaneous polarization**, known as **domains**, emerge in an insulator. Ferroelectrics, for example, have a domain structure.

In the normal state, domains in ferroelectrics have a random orientation due to the thermal motion. But when a ferroelectric is introduced into an external electrostatic field, domains in the crystal acquire a preferred orientation, and the substance is polarized. Since the effect of polarization (abruptly) involves the whole regions, the internal electric field in ferroelectrics may attain considerable values.

Another type of insulators is presented by **electrets**. These are the materials that are polarized artificially in the liquid state. Then electrets are transformed to the solid state, and the polarizing external electrostatic field is removed. As a result, a residual polarization emerges in the body. Such a state is just an electric analogue of permanent magnets.

These exotic states of a substance are not numerous. In actual practice, an external electrostatic field independently polarizes atomic or molecular dipole moments. A partial polarization sets in since the thermal perturbation is a competing process. If an insulator does not have a dipole structure, a sufficiently strong external field may "deform" electron shells of atoms or molecules and move apart the "centres of positive and negative charges", i.e. dipole states are induced and polarized simultaneously.

Thus, a typical feature of the behaviour of an insulator in an external electrostatic field is the emergence of the polarization state and accordingly the internal electric field. It suppresses the external field, and the electric field of strength

$$\mathbf{E}_{\text{int}} = \mathbf{E}_{\text{ext}} - 4\pi\mathbf{P} \qquad (7.5.9)$$

is established in the insulator, where $\mathbf{P}$ is the dipole moment of a unit volume of the insulator.

Let us consider the main quantitative relations for insulators in electrostatic fields by using a gas as an example. In the absence of a field, individual atoms of the gas, which are in the ground quantum state, have a spherical symmetry of the averaged electron distribution in the space surrounding a nucleus. This means that the centres of positive and negative charges in such atomic states coincide. The application of a constant external electric field creates an axial symmetry in atomic states, which is accompanied by the emergence of the electric dipole moments of the atoms. For the sake of simplicity, let us consider atomic hydrogen. Two forces act upon a deformed hydrogen atom: the force $e\mathbf{E}$ exerted by an

external field on a charge, and the "elastic" restoring force $m\omega_0^2\,\Delta r$, where m is the mass of the charge, Δr is the displacement of the centres of the positive and negative charges in the atom relative to each other, and ω_0 is a quantity having dimensions s^{-1}. It should be noted that any elastic restoring force can be represented in this form (see also Sec. 5.10.3). As regards ω_0, in the case under consideration, it can be estimated from the ratio $\mathscr{E}/\hbar$, where $\mathscr{E}$ is the energy of the ground state of the atom (see, for example, Eq. (4.7.1)).

Therefore, we can write

$$eE = m\omega_0^2\,\Delta r, \tag{7.5.10}$$

where $\omega_0 = \mathscr{E}/\hbar$ and $\mathscr{E} = (1/2)\,(me^4)/\hbar^2$ (see Eq. (4.5.8)).

Relation (7.5.10) allows us to determine the emerging dipole moment of the hydrogen atom. Indeed, the dipole moment is given by

$$\mathbf{d} = e\,\Delta\mathbf{r} \tag{7.5.11}$$

(see the explanation to Eq. (3.4.17)). The displacement can be written as

$$\Delta r = \frac{eE}{m\omega_0^2} = \frac{eE}{m\left(\frac{\mathscr{E}}{\hbar}\right)^2} = \frac{eE}{m\left(\frac{1}{2}\,\frac{me^4}{\hbar^3}\right)^2} = 4\,\frac{r_a^3}{e}\,E, \tag{7.5.12}$$

where r_a represents the "size" of the hydrogen atom (see Eq. (4.5.7)). Substituting (7.5.12) into (7.5.11), we obtain

$$\mathbf{d} = 4r_a^3\mathbf{E}. \tag{7.5.13}$$

The dipole moment of a unit volume of the gas is

$$\mathbf{P} = N\mathbf{d} = 4Nr_a^3\mathbf{E}_{\text{int}}, \tag{7.5.14}$$

where N is the number of atoms in a unit volume. As regards the electric field appearing in (7.5.14), it should be noted that unlike for an isolated atom in an external field (see Eq. (7.5.13)), for the aggregate of atoms we must consider the field which is the superposition of the external and the mean internal field of the atoms being polarized. Let us write Eq. (7.5.9) in the form

$$\mathbf{E}_{\text{ext}} = \mathbf{E}_{\text{int}} + 4\pi\mathbf{P} \tag{7.5.15}$$

and substitute the obtained expression for the dipole moment **P** of the medium into this equation. This gives

$$\mathbf{E}_{ext} = \mathbf{E}_{int} + 4\pi \cdot 4Nr_a^3 \mathbf{E}_{int} = (1 + 16\pi r_a^3 N)\, \mathbf{E}_{int}. \qquad (7.5.16)$$

The quantity

$$\alpha \approx 16\pi r_a^3 \qquad (7.5.17)$$

is referred to as the **polarizability** of the atom, and the quantity

$$\varepsilon = 1 + \alpha N \qquad (7.5.18)$$

is known as the **permittivity** of the medium. For vacuum ($N = 0$),

$$\varepsilon = 1. \qquad (7.5.19)$$

Let us determine the numerical value of ε for atomic hydrogen under normal conditions. For hydrogen, $r_a = 0.528 \times 10^{-10}$ m (see Sec. 4.5.1), and the number of atoms in a cubic metre of the gas is $N = 2.69 \times 10^{25}$ (see Sec. 5.8.3). As a result, according to Eq. (7.5.18), we obtain

$$\varepsilon = 1 + 16\pi\, (0.528 \times 10^{-10})^3\, (2.69 \times 10^{25}) = 1.00020.$$

The experimental value of permittivity for molecular hydrogen is

$$\varepsilon_{exp} = 1.00026.$$

The agreement between the theoretical and experimental values of ε for hydrogen is quite satisfactory. Indeed, firstly, calculations were made for atomic hydrogen, while the experiment was made with molecular hydrogen (a small difference should exist). Secondly, the calculations themselves were made in the framework of semiclassical model and not with the help of the consistent quantum-mechanical theory which gives the value of α differing from Eq. (7.5.17) in the numerical factor (18π instead of 16π).

Let us consider in greater detail the polarization of a gas with molecules having an intrinsic electric dipole moment **d**. In the absence of an external electrostatic field, all molecular dipoles **d** are oriented randomly because of thermal motion. The application of a field **E** introduces a certain preferred orientation of dipoles along the field. According to

Eq. (3.4.18), the energy of a dipole **d** in the external field **E** is

$$U = -\mathbf{d}\cdot\mathbf{E}. \tag{7.5.20}$$

The energy of thermal motion associated with an individual molecular dipole is of the order of kT. Thermal motion produces a disorienting effect on the dipole. Therefore, the dipole moment **P** can be presented in the form

$$|\mathbf{P}| \sim \left(\frac{|\mathbf{d}|\,|\mathbf{E}|}{kT}\right) N\,|\mathbf{d}|. \tag{7.5.21}$$

This formula shows that the maximum polarization of the gas could take place at as low temperatures as possible and for considerably high values of electric field strength, i.e. when $|\mathbf{d}|\,|\mathbf{E}| \approx kT$.

The exact expression for the orientation dipole moment **P** differs from (7.5.21) in the factor 1/3, i.e.

$$P = \frac{Nd^2E}{3kT}. \tag{7.5.22}$$

Accordingly, the orientation dipole permittivity of a gas with dipole molecules will be

$$\varepsilon - 1 = \frac{P}{E} = \frac{Nd^2}{3kT}. \tag{7.5.23}$$

The permittivity (to be more precise, $\varepsilon - 1$) of such a gas is proportional to the number density N of particles and inversely proportional to temperature.

While analyzing the orientation polarizability, we stated that molecular dipoles are predominantly oriented along the field. However, they can also be oriented against the field. If we proceed from expression (7.5.20) for the energy of a dipole in an electric field,

$$U = -\,|\mathbf{d}|\,|\mathbf{E}|\cos\,(\widehat{\mathbf{dE}}), \tag{7.5.24}$$

the angle $(\widehat{\mathbf{dE}}) = 0$ when the dipole moment is parallel to the electric field strength, and hence $\cos\,(\widehat{\mathbf{dE}}) = 1$. If, however, they have antiparallel directions, $(\widehat{\mathbf{dE}}) = \pi$, and hence $\cos\,(\widehat{\mathbf{dE}}) = -1$. Therefore, the energy of a dipole changes with the orientation angle from the maximum (pos-

itive) value $U_{max} = |\mathbf{d}|\,|\mathbf{E}|$ to the minimum (negative) value $U_{min} = -|\mathbf{d}|\,|\mathbf{E}|$. Since the energy of any system tends to assume the minimum possible value for given conditions, the most probable dipole orientation will be that corresponding to U_{min}, i.e. in the direction of the electric field.

In order to determine the mean number of molecular dipoles oriented along the field and against it, we must resort to the methods of statistical physics (see Chap. 5). In thermal equilibrium, the relative number of molecules with potential energy U is proportional to

$$e^{-U/(kT)} \tag{7.5.25}$$

(see Eq. (5.6.3)). Substituting (7.5.24) into (7.5.25), we obtain

$$n(\theta) = n_0 e^{+|\mathbf{d}|\,|\mathbf{E}|\cos\theta/(kT)}. \tag{7.5.26}$$

In this formula, we have introduced the notation for the angle $\theta \equiv (\widehat{\mathbf{dE}})$. The quantity $n(\theta)$ is the mean number of molecular dipoles contained in a unit solid angle about the direction θ, and n_0 is the normalization constant. Since the expression in the parentheses of (7.5.21) is small under normal physical conditions, the exponent in (7.5.26) can be transformed in accordance with (4.4.9). As a result, (7.5.26) can be written in the form

$$n(\theta) \approx n_0 \left(1 + \frac{|\mathbf{d}|\,|\mathbf{E}|\cos\theta}{kT}\right). \tag{7.5.27}$$

This relation shows that a larger number of dipoles are oriented along the field ($\cos\theta = 1$) than against the field ($\cos\theta = -1$). The mean dipole moment of the gas is due to this difference.

7.6. A Substance in a Magnetic Field

7.6.1. Diamagnetic Effect. The interaction of a magnetic field with a substance is a relativistic phenomenon. Since the ratio v/c for atomic electrons is small, the effect of the magnetic field on the substance is weak (of the order of v^2/c^2). However, there are states of a substance in which the response to an external magnetic effect is significant and peculiar.

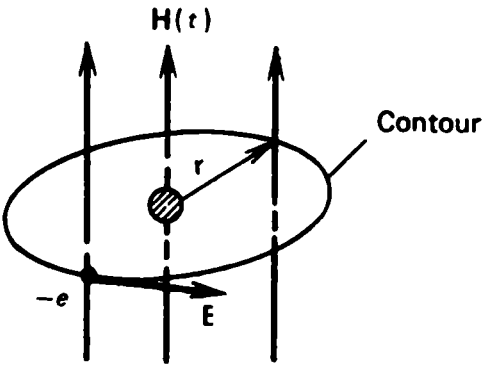


FIG. 7.7

The nature of the magnetic properties of a substance can be completely explained only by the quantum theory. It is well known that the total mechanical angular momentum (orbital and spin) of all electrons in atoms can either differ from zero or be equal to zero. Since the mechanical angular momentum is associated with a magnetic moment, the atoms of different elements either have a magnetic moment or do not have any. Let us consider the response of atoms that have no "intrinsic" magnetic moment to the application of an external magnetic field of strength **H**. Let us suppose that the magnetic field in the vicinity of an atom slowly increases (Fig. 7.7). In this case, a varying magnetic flux passes through the contour embracing the region of motion of the electrons. According to Eq. (3.3.6), this flux generates a vortex electric field. Therefore,

$$\langle E_\tau \rangle 2\pi r = -\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}(\langle H_n \rangle \pi r^2). \tag{7.6.1}$$

For the mean tangential component of the electric field strength on the contour, we obtain

$$\langle E_\tau \rangle = -\frac{1}{c}\frac{r}{2}\frac{\mathrm{d}\langle H_n \rangle}{\mathrm{d}t}. \tag{7.6.2}$$

Let us determine the change in the mechanical angular momentum of atomic electrons under the effect of the field $\langle E_\tau \rangle$. For this purpose, we generalize the classical equation of motion (2.3.1)

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{F} \tag{7.6.3}$$

to the case of rotational motion. Premultiplying both sides of Eq. (7.6.3) by the vector **r**, we obtain

$$\frac{\mathrm{d}[\mathbf{rp}]}{\mathrm{d}t} = [\mathbf{rF}]. \tag{7.6.4}$$

Here, [rp] is the mechanical angular momentum **L** (see Eq. (2.6.26)), and the vector product [rF] is known as the **moment of force**. As a result, Eq. (7.6.4) assumes the form

$$\frac{d\mathbf{L}}{dt} = [\mathbf{rF}]. \tag{7.6.5}$$

Since in the case under consideration, $F = -eE$ (we assume the geometry of Fig. 7.7), Eq. (7.6.5) will be written in the form

$$\frac{dL}{dt} = \frac{e}{2c}\frac{dH}{dt}\sum r^2, \tag{7.6.6}$$

where the summation is carried out over all the atomic electrons. In this equation, we have used Eq. (7.6.2). Integrating (7.6.6) with respect to time beginning with the zero magnetic field over the time interval of variation of the field **H**, we obtain

$$\Delta L = \frac{e}{2c} H \sum r^2. \tag{7.6.7}$$

This equation describes the variation of the mechanical angular momentum of atomic electrons during the time of application of the magnetic field.

Since we consider the atom in which the intrinsic mechanical angular momentum is zero, the emergence in it of the angular momentum $\Delta\mathbf{L}$ (over the period of application of the magnetic field or in a strongly nonuniform magnetic field) is accompanied by the emergence of the corresponding magnetic moment $\mathbf{p}_m$ (see Secs. 3.5.3 and 4.5.2). According to Eq. (4.5.13), for the induced magnetic moment of the atom we can write

$$\Delta p_m = -\frac{e}{2mc}\Delta L = -\frac{e^2}{4mc^2} H \sum r^2. \tag{7.6.8}$$

The minus sign in this expression indicates that the magnetic moment is antiparallel to the external magnetic field. For this reason, this magnetic moment is referred to as a **diamagnetic moment**.

Obviously, formula (7.6.8) must contain $\sum \langle r^2 \rangle$ instead of $\sum r^2$. This sum can be found only in the framework of quantum mechanics (this can be done, for example, with the help of Fig. 4.13*b*).

The manifestation of diamagnetic effects on the macroscopic scale is known as **diamagnetism** of a substance. Typical diamagnets (like bismuth) placed in a highly nonuniform magnetic field are pulled out from the region of the strong field.

7.6.2. Paramagnets. Orientation Magnetization. Let us consider atoms possessing an intrinsic mechanical angular momentum. It is well known that the orbital angular momentum, and hence the magnetic moment of an atomic electron, are subjected to space quantization (see Sec. 4.5.2). This means that one of the projections of the magnetic moment on the direction of a magnetic field assumes only a discrete number of values, its two other projections being indeterminate. This situation resembles the precession of a gyroscope about the vertical direction.

The atomic magnetic moment $\mathbf{p}_{ma}$ in a magnetic field of strength $\mathbf{H}$ has the energy

$$U = -\mathbf{p}_{ma}\mathbf{H} \tag{7.6.9}$$

(see Eq. (4.5.14)). The electron charge appearing in the expression for $\mathbf{p}_{ma}$ is denoted here by $-e$. Expression (7.6.9) is similar to (7.5.20). This means that the quantitative theory of orientation magnetization of a substance can be constructed by analogy with the theory of orientation polarizability. Let us introduce the concept of **magnetization** of a unit volume of a substance:

$$\mathbf{J} = \sum_i \mathbf{p}_{ma_i} = N \langle \mathbf{p}_{ma} \rangle, \tag{7.6.10}$$

where N is the number density of atoms, and $\langle \mathbf{p}_{ma} \rangle$ is the mean magnetic moment of an atom. Then, according to Eq. (7.5.22), we can write

$$\mathbf{J} = \frac{N \langle (\mathbf{p}_{ma}) \rangle^2 \mathbf{H}}{3kT}. \tag{7.6.11}$$

Thus, *the magnetization of a substance is proportional to the external magnetic field strength.* Such a behaviour of the substance is known as **paramagnetism**. This property is manifested most clearly at low temperatures.

The limits of applicability of formula (7.6.11) are set by the smallness of the ratio $|\mathbf{p}_{ma}|\,|\mathbf{H}|/(kT)$. In order to

estimate the temperature at which the **magnetic saturation** of a substance can be observed, i.e. when all atomic magnetic moments are oriented along the field, we shall use estimate (4.5.14). As a result, we find that the saturation temperature $T \sim 10$ K (for fields $H \sim 10^7$ A/m).

There exists an interesting method of obtaining ultralow temperatures by adiabatic demagnetization. A paramagnetic substance is cooled in a strong magnetic field to the liquid helium temperatures (1-2 K), i.e. is brought to the state of magnetic saturation. Then the substance is thermally insulated, and the magnetic field is slowly removed. The temperature of the substance drops thereby to a few thousandths of kelvin.

The essence of the process of adiabatic demagnetization lies in the following. Under magnetic saturation at $T \to 0$, the substance is in its lowest quantum state, i.e. $\Delta\Gamma = 1$ and the entropy $S = 0$ for it (see Sec. 5.4). As the magnetic field is gradually removed, the entropy must increase in view of irreversibility of the process. It means that the number $\Delta\Gamma$ of microscopic ways of realization of the given state of the substance increases. In other words, magnetic transitions to higher energy states will take place. The process of randomization of atomic magnetic moments occurring in a decreasing magnetic field requires some energy expenditures. Under the conditions of thermal insulation of the substance, the energy is borrowed from the thermal motion of the lattice.

Let us briefly consider the elements whose atoms have nonzero magnetic moments. All the atoms with an odd number of electrons have a constant magnetic moment whose value is close to

$$p_{\mathrm{ma}} = \frac{e\hbar}{2mc}. \tag{7.6.12}$$

Substituting the numerical values expressed in the CGS units and converting the result into the SI units, we obtain

$$p_{\mathrm{ma}} \simeq 9 \times 10^{-24} \text{ J/T}.$$

Comparing (7.6.12) with (4.6.3) and (4.6.4), we can easily see that the atomic magnetic moment is due to the noncompensated spin of one ("odd") electron in an atom. Then it becomes clear that the molecules of most substances

cannot be "atomic magnets". Indeed, molecular compounds are formed in such a way that the valence electrons of the atoms which supplement each other and have opposite spins are paired.

Paramagnets introduced in a highly nonuniform magnetic field are pulled into the region of the strong field, i.e. their "macroscopic behaviour" is opposite to that of diamagnets under similar external conditions. It should be noted that since the diamagnetic effect is of universal nature, it is also present in paramagnets. However, an additional induced diamagnetic moment of an atom is as a rule smaller than the constant magnetic moment.

7.6.3. Spontaneous Magnetization. Ferromagnetism. Diamagnetic and paramagnetic effects are rather weak. There exist, however, "composite materials" whose magnetic properties are very strong. As a rule, such materials contain atoms with an unfilled inner electron shell. These are the atoms of transition elements like iron, cobalt, and nickel.

In the absence of an external magnetic field, small iron or nickel crystals exhibit the state of complete spontaneous magnetization of a substance at a temperature below a certain value. This is **spontaneous magnetization** which requires an explanation. A possible mechanism of this phenomenon is as follows. In a nickel atom, for example, a constant atomic magnetic moment is due to the spin of an electron from the unfilled inner shell. It turns out that these spins in a nickel lattice align themselves in a certain direction. Let us suppose that the parallelism of the spins is ensured by their direct magnetic interaction. The energy of the spin-spin interaction can be estimated with the help of the formula

$$U_{ss} \sim \frac{(\mathbf{p}_{ms})^2}{r_a^3} \tag{7.6.13}$$

(see Eq. (4.6.5)). Since the interatomic distance in the lattice is of the order of r_a, using formula (4.5.7), we can write

$$U_{ss} \sim \frac{1}{c^2}\left(\frac{me^8}{\hbar^4}\right). \tag{7.6.14}$$

Substituting the numerical values expressed in the CGS units and converting the result into the SI units, we obtain

$U_{ss} \sim 5 \times 10^{-22}$ J.

What is the temperature corresponding to this value of energy? It can be found from the relation

$U_{ss} \sim kT$,

i.e. $T \sim 10$ K.

This means that the thermal perturbation of the lattice does not allow spins to align themselves in one direction. Thus, the assumption about the role of the magnetic interaction of spins in the formation of the state of spontaneous magnetization is incorrect.

Another possible mechanism that explains the alignment of the spins is the indirect quantum-mechanical effect of "magnetic" electrons on one another. This effect is due to the fact that electrons are Fermi particles. Electrons can be close neighbours when their spins are antiparallel. Let us suppose that there exists such a proximity between inner "magnetic" electrons and conduction (free) electrons. Then we shall have a situation depicted in Fig. 7.8. Conduction electrons play the role of an intermediary between the "magnetic" electrons of neighbouring atoms. The spin of a "magnetic" electron from the unfilled inner shell of an atom is antiparallel to the spin of a free electron. Such an orientation of the latter spin ensures the parallelism of the magnetic moments of the atoms themselves. The intensity of this indirect quantum-mechanical effect of atomic magnetic moments on one another is 10^4 times as strong as their

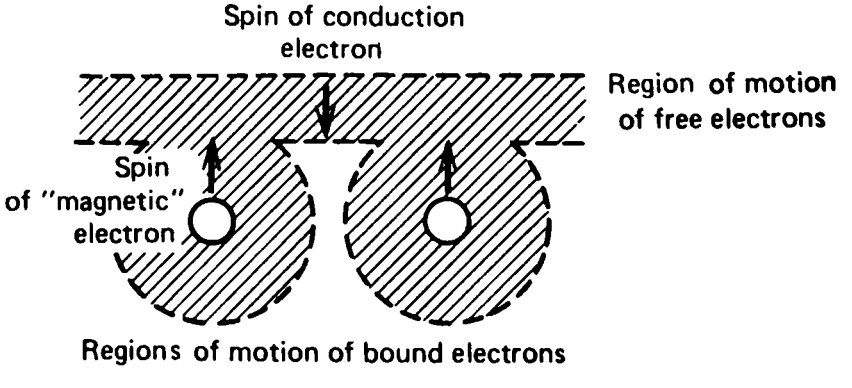

Spin of conduction electron
Region of motion of free electrons
Spin of "magnetic" electron
Regions of motion of bound electrons

FIG. 7.8

direct magnetic interaction. This estimate follows from the fact that

$$U_{ss} \sim 10^{-4} E_a \sim 10^{-4} \varepsilon_F, \tag{7.6.15}$$

where E_a is the energy of an atomic electron (see Eq. (4.5.8)). The energies ε_F of free (intermediary) electrons are comparable with E_a (see Eq. (5.10.24)). Thus, we have considered a feasible mechanism of spontaneous magnetization of a substance. At any rate, this mechanism is adopted at present by theoretical physicists.

The state of spontaneous magnetization of a substance is called **ferromagnetism**. As was mentioned above, only very small crystals of ferromagnets (whose size is normally less than 10^{-6} m) exhibit the complete magnetic saturation in the absence of external magnetic fields. In large ferromagnetic samples, there is no complete magnetization in the absence of an external field. This is due to the fact that the entire aggregate of atoms in such samples is split into magnetic regions, viz. domains. Within a domain, the complete magnetization does exist, but different domains have different directions of magnetization relative to one another.

Small crystals of ferromagnets contain a single domain. Why is a large number of domains formed in, say, a large iron single crystal? The answer is because it is advantageous from the point of view of thermodynamics. The "free energy" of any physical system tends to attain the minimum value. If a large iron single crystal contains only one domain, the energy of the corresponding magnetic field outside the crystal attains the maximum possible value. A magnetic splitting of the single crystal into two regions with antiparallel magnetizations sharply reduces the "field volume" (and hence its energy) outside the crystal (Fig. 7.9).

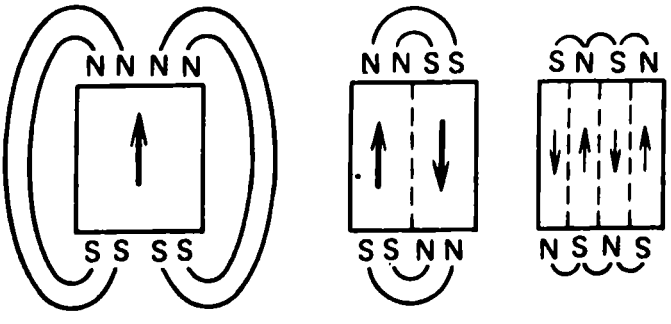


FIG. 7.9

Only a sufficiently strong external magnetic field $\mathbf{H}_{ext}$ can orient all the domains of a large ferromagnetic sample in the same direction. Then the field in the bulk of the sample is given by

$$\mathbf{H}_{int} = \mathbf{H}_{ext} + 4\pi\mathbf{J}, \tag{7.6.16}$$

where $\mathbf{J}$ is the magnetization of the substance (see Eq. (7.6.10)).

In order to describe the magnetic properties of a substance, we can write the approximate relation

$$\mathbf{H}_{int} = \mu\mathbf{H}_{ext}, \tag{7.6.17}$$

where μ is the **absolute permeability** of the substance. For diamagnets, $\mu \lesssim 1$ and differs from unity by 10^{-5}-10^{-6}. For paramagnets, $\mu \gtrsim 1$ and differs from unity by 10^{-4}-10^{-5}. For ferromagnets, $\mu \sim 10^3$ (iron) and $\mu \sim 10^6$ for supermalloy.

It should be noted that for all substances, the permittivity $\varepsilon > 1$, and for ferroelectrics (barium titanate), $\varepsilon \sim 10^5$.

The data presented above indicate that the response of most solids and liquids to the direct effect of a magnetic field is 10^{-4} times as strong as to the effect of an electric field. The same conclusion follows from the fact that the energy of interaction of two magnetic moments amounts to only 10^{-4} of the charge-charge interaction energy (see Eq. (7.6.15)). This conclusion can also be drawn from the remark made at the beginning of Sec. 7.6.1. Indeed, for atomic electrons, $v/c \sim 10^{-2}$, and the effect of a magnetic field on a substance is proportional to v^2/c^2, i.e. $(10^{-2})^2 \sim 10^{-4}$.

A special position is occupied by ferromagnets and ferroelectrics: the former possess spontaneous magnetization and the latter, spontaneous polarization. In such states of a substance, external magnetic and electric effects on macroscopic bodies "hold their own".

7.6.4. Magnetic Properties of Superconductors. Quantization of Large-Scale Magnetic Flux. The nature of superconducting transition in metals was described in Sec. 7.1. The absence of electrical resistance means that an external magnetic field cannot penetrate the bulk of a superconductor, i.e.

$$\mathbf{H}_{int} = 0. \tag{7.6.18}$$

This result follows directly from the fundamental field equation (3.3.6) since any penetration of the magnetic field into

a medium is associated with the process of its stabilization, i.e. with its change in time. But according to Eq. (3.3.6), this leads to the emergence of a vortex electric field, and hence of an indefinitely strong current (since the resistance is zero). We have arrived at an absurd result which shows that the assumption on the possible emergence of a magnetic field in the bulk of a superconductor is erroneous.

Expression (7.6.18) obtained above holds irrespective of the previous history of the establishment of the superconducting state in a metal. For example, if a sample of a normal metal is in a magnetic field and its temperature decreases, the magnetic field will be "pushed out" from the bulk of the sample at the temperature corresponding to the superconducting transition.

What is the configuration of the magnetic field surrounding a superconductor? The answer to this question can be obtained with the help of another fundamental field equation (3.3.4). Since the magnetic field is zero in the bulk of the superconductor, the component normal to the boundary with a medium is naturally also zero. And since the general relation (3.3.4) must be valid, it means that the component of the external magnetic field, which is normal to the surface, must also be zero. In other words, *the external magnetic field as if flows past the superconducting sample.* Only the tangential components of the field differ from zero near the surface.

Let us consider a superconducting ring. According to the general field equation (3.3.5), no stationary current can flow in the bulk of the superconductor since $\mathbf{H}_{\text{int}} = 0$. Thus, *only a surface electric current can exist in the superconducting ring.*

In particular, the last conclusion means that the magnetic field must penetrate at least a narrow surface layer of a superconductor. Measurements show that the penetration depth of magnetic field into a superconductor is of the order of 10^{-7} m. About a thousand atomic layers of the substance fit into such a depth.

Thus, let a superconduction current exist in the ring. The current state in a superconductor is known to be the lowest collective quantum state for electron pairs which are bosons (see Sec. 7.1). The current of doubly charged carriers (electron pairs) is macroscopic, the current assuming only a number of discrete values. The large-scale magnetic flux

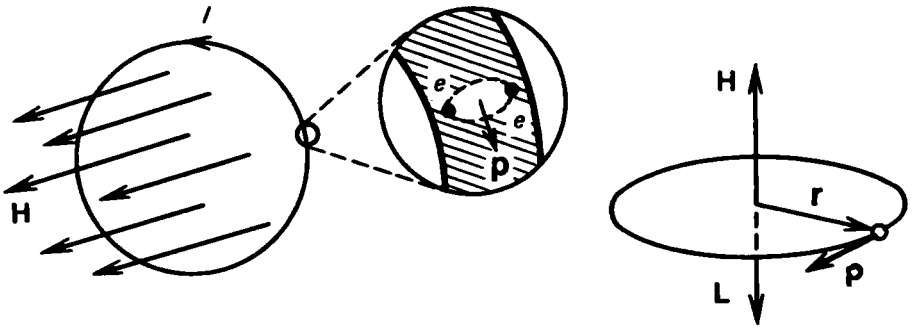


FIG. 7.10

(through the superconducting ring) produced by this current is quantized.

In order to describe the above phenomena quantitatively, we shall use, for the sake of simplicity, the semiclassical approach combining the features of the classical method (using mechanics and electrodynamics) and ideas of quantization. We proceed from the expression for the angular momentum of an electron pair (see Eq. (2.6.26))

$$\mathbf{L} = [\mathbf{rp}], \tag{7.6.19}$$

where $\mathbf{r}$ is the radius vector of the centre of mass of the pair, and $\mathbf{p}$ is the momentum of the pair transfer (Fig. 7.10). Let us express the momentum $\mathbf{p}$ in terms of electrodynamic parameters. The magnetic component of the Lorentz force (3.2.3)

$$\mathbf{F} = \frac{Q}{c}[\mathbf{vH}]$$

acting on the electron pair ($Q = 2e$) in a constant uniform magnetic field $\mathbf{H}$ produced by the superconduction current itself can be transformed as follows:

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{Q}{2c}\left[\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\mathbf{H}\right] = \frac{Q}{2c}\frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{rH}].$$

Omitting the operation $\mathrm{d}/\mathrm{d}t$ on the left- and right-hand sides, we obtain the following expression for the momentum of the electron pair:

$$\mathbf{p} = \frac{Q}{2c}[\mathbf{rH}]. \tag{7.6.20}$$

Substituting (7.6.20) into (7.6.19), we get

$$\mathbf{L} = \frac{Q}{2c}[\mathbf{r}[\mathbf{rH}]] = \frac{Q}{2c}\mathbf{r}(\mathbf{rH}) - \frac{Q}{2c}\mathbf{H}(\mathbf{rr}) = -\frac{Q}{2c}\mathbf{H}r^2. \tag{7.6.21}$$

Here, we have used the following rule for the triple vector product:

$$[\mathbf{a}[\mathbf{bc}]] = \mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab}) \tag{7.6.22}$$

and the fact that $(\mathbf{rH}) = 0$ since $\mathbf{r} \perp \mathbf{H}$.

Let us write (7.6.21) in the form

$$|\mathbf{L}| = \frac{Q}{2\pi c} H\pi r^2 = \frac{Q}{2\pi c}\Phi, \tag{7.6.23}$$

where Φ is the magnetic flux across the surface stretched on the superconducting current-carrying ring (see Eq. (3.3.1)). Since the angular momentum of any quantum system (macroscopic in the given case) can assume only a discrete set of values, expression (7.6.23) becomes

$$n\hbar = \frac{Q}{2\pi c}\Phi_n. \tag{7.6.24}$$

Here, Φ_n is the quantized magnetic flux, and $n = \pm 1, \pm 2, \ldots$. Thus, we obtain the following expression for a quantum of the large-scale magnetic flux:

$$\Phi_1 = \pi\frac{\hbar c}{e}. \tag{7.6.25}$$

Substituting the numerical values expressed in the CGS units and converting the result into the SI units, we obtain

$$\Phi_1 \simeq 2 \times 10^{-15}\ \text{Wb}.$$

Such a flux can easily be created in a small ring of 1-mm diameter.

Of course, such an analysis of current in the superconducting ring is far from being consistent. It would be better to use the macroscopic wave function of a superconducting "condensate" and all the conditions imposed on this function. Such an approach would also lead to (7.6.25).

A superconducting current-carrying ring resembles a hydrogen atom. In this connection, it should be noted that the

sizes of atoms in excited states become much larger. Indeed, writing expression (4.5.9) in the form

$$\frac{e^2}{r_n} = \frac{e^2}{r_a n^2},$$

we obtain

$$r_n = r_a n^2.$$

For example, hydrogen atoms excited to 100-150th state exist in the free space. For $n = 100$, the radius of a hydrogen atom increases by a factor of 10^4, i.e. its size is about 10^{-6} m.

7.7. Alternating Currents and Electromagnetic Waves in a Medium. Optical Properties of Media

7.7.1. A.C. Fields and a Substance. The response of a substance to the external effect of electric and magnetic fields is determined not only by the state of the substance (say, its temperature) but also by the frequency of the external fields (their dependence on time). When the frequency ω of an external electromagnetic field is high enough, the values of permittivity ε and permeability μ of the substance tend to their values in vacuum. At frequencies ω comparable with natural (characteristic) frequency ω_0 of the substance itself, like the frequencies due to oscillations of molecular electric dipole or magnetic atomic moments, the frequency dependences $\varepsilon(\omega)$ and $\mu(\omega)$ become very strong.

Let us first consider alternating fields for which the wavelength $\lambda \sim c/\omega$ corresponding to the frequency ω of the field is large in comparison with the size l of a body:

$$\lambda \gg l$$

or

$$\omega \ll \frac{c}{l}. \qquad (7.7.1)$$

This condition indicates that we neglect the lag due to the finite velocity of propagation of electromagnetic perturbations. As long as the frequency of an alternating current satisfies condition (7.7.1), the instantaneous value of the

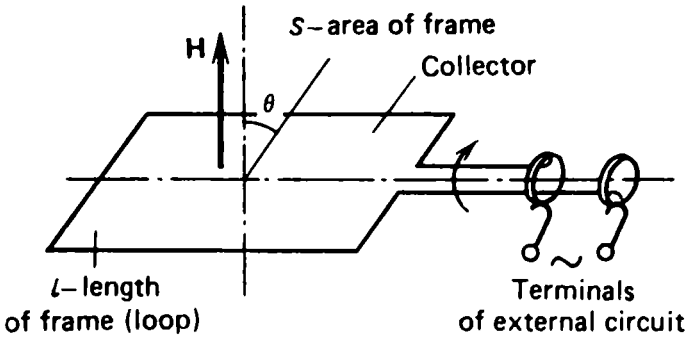


FIG. 7.11

current $I(t)$ in the circuit is determined by the value of the emf $\mathscr{E}(t)$ at this instant from the relation

$$I(t)=\frac{1}{R}\mathscr{E}(t) \tag{7.7.2}$$

(see Eq. (7.3.10)).

7.7.2. Induced EMF. A practically convenient method of obtaining an alternating emf $\mathscr{E}_1$ is as follows. A rotating conducting loop (frame) of length l is placed into an external magnetic field **H**. According to Eq. (3.3.6), we then have

$$\mathscr{E}_1 \equiv \langle E_\tau \rangle l = -\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}(\langle H_n \rangle S) \equiv -\frac{1}{c}\frac{\mathrm{d}\Phi}{\mathrm{d}t}. \tag{7.7.3}$$

Here, the emf emerges due to a circulating electric field **E** which does the work on the displacement of charges along the loop l.

Let us introduce a conducting frame (loop) into a uniform magnetic field **H** (Fig. 7.11). Suppose that the frame rotates with an angular velocity ω. The position of the frame at any instant t will be specified by the angle $\theta = \omega t + \alpha$, where the constant α determines the orientation of the frame at $t = 0$. The magnetic flux Φ across the plane of the frame is equal to the product of the normal component of the magnetic field $H_n = H \sin\theta$ by the area S of the surface embraced by the frame, i.e.

$$\Phi(t) = SH \sin(\omega t + \alpha). \tag{7.7.4}$$

According to (7.7.3), for the induced emf we obtain

$$\mathscr{E}_1 = -\frac{1}{c}\frac{d\Phi}{dt} = -\frac{SH\omega}{c}\cos(\omega t + \alpha), \tag{7.7.5}$$

or

$$\mathscr{E}_1 = \mathscr{E}_0 \cos \omega t. \tag{7.7.6}$$

If an emf of the type (7.7.6) is introduced into an electric circuit, phenomena which can be referred to as **self-induction** may take place in individual regions of the circuit. As a matter of fact, the periodic emf (7.7.6) will cause a periodic current $I(t)$ (see Eq. (7.7.2)). And the current is accompanied by the emergence of an intrinsic magnetic field (see, for example, Eq. (3.5.2)) whose flux is

$$\Phi_{\text{intr}}(t) = \frac{1}{c} LI(t), \tag{7.7.7}$$

where L is a certain coefficient having the dimensions of length and depending on the shape and size of the conductor. Moreover, L explicitly contains the magnetic characteristics μ of the medium surrounding the conductor. The alternating intrinsic magnetic flux $\Phi_{\text{intr}}(t)$ will induce an intrinsic emf

$$\mathscr{E}_{s.i} = -\frac{1}{c}\frac{d}{dt}\Phi_{\text{intr}}$$

$$= -\frac{1}{c}\frac{d}{dt}\left(\frac{1}{c}LI\right) = -\frac{1}{c^2}L\frac{dI}{dt}. \tag{7.7.8}$$

This emf is known as the **self-induced emf**, while the coefficient L is referred to as the **self-inductance**. The value of L is especially large for conducting coils with iron cores.

7.7.3. A.C. Circuits. Solutions of Differential Equations. Let us make up a circuit from a resistor of resistance R and inductance L (this is the notation for an element of the circuit containing a coil with self-inductance L). We connect the circuit to an external periodic emf $\mathscr{E}_1$ (Fig. 7.12) and analyze the time dependence $I(t)$ of the current in the circuit. The equation for this circuit will be based on relation

(7.3.10). For the case under consideration, it has the form

$$IR = \mathcal{E}_i + \mathcal{E}_{s.i}, \tag{7.7.9}$$

or

$$\frac{1}{c^2} L \frac{dI}{dt} + RI = \mathcal{E}_i. \tag{7.7.10}$$

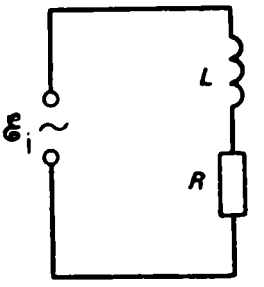


FIG. 7.12

How can we find the solution $I(t)$ of the differential equation (7.7.10)? The form of this solution can be predicted qualitatively beforehand. Indeed, considering only the steady-sitate conditions in the circuit and neglecting all trans ent processes, we can state that the current must perform periodic forced oscillations, but with a certain phase shift relative to the external periodic emf. Thus, let the solution of the equation

$$\frac{1}{c^2} L \frac{dI}{dt} + RI = \mathcal{E}_0 \cos \omega t \tag{7.7.11}$$

have the form

$$I(t) = I_0 \cos(\omega t + \varphi), \tag{7.7.12}$$

in which the unknown constants I_0 and φ must be specified.

In order to simplify the form of the original equation and to use practical units convenient for electrical-engineering calculations (ampere, volt, ohm, etc.), we go over to SI units. For this purpose, we use the conversion factors from CGS to SI (see Appendix II, Table II.1):

$$c^2 \to \frac{1}{\varepsilon_0 \mu_0}, \quad L \to \frac{4\pi}{\mu_0} L, \quad R \to 4\pi\varepsilon_0 R,$$

$$I \to \frac{1}{\sqrt{4\pi\varepsilon_0}} I, \quad \mathcal{E}_0 \to \sqrt{4\pi\varepsilon_0}\, \mathcal{E}_0.$$

As a result, equation (7.7.11) assumes the form

$$L \frac{dI}{dt} + RI = \mathcal{E}_0 \cos \omega t, \tag{7.7.11a}$$

i.e. the constant c^2 falls out of it.

Thus, we must determine the constants I_0 and φ in solution (7.7.12). For this purpose, we substitute (7.7.12) into (7.7.11a):

$$-LI_0\omega \sin(\omega t + \varphi) + RI_0 \cos(\omega t + \varphi) = \mathcal{E}_0 \cos \omega t.$$

Using the well-known formulas for sine and cosine of the sum of two angles, we obtain

$$-LI_0\omega(\sin \omega t \cos \varphi + \cos \omega t \sin \varphi)$$
$$+RI_0(\cos \omega t \cos \varphi - \sin \omega t \sin \varphi) = \mathcal{E}_0 \cos \omega t.$$

Transferring all the terms of this equation to the left-hand side, we write it in the form

$$A \cos \omega t + B \sin \omega t = 0.$$

Since this equality is satisfied if A and B are simultaneously equal to zero, we obtain

$$-LI_0\omega \sin \varphi + RI_0 \cos \varphi - \mathcal{E}_0 = 0, \tag{7.7.13}$$
$$-LI_0\omega \cos \varphi - RI_0 \sin \varphi = 0. \tag{7.7.14}$$

Using (7.7.14), we obtain the following expression for the phase shift between current and emf:

$$\tan \varphi = -\frac{L\omega}{R} \tag{7.7.15}$$

or

$$\cos \varphi = \frac{1}{\sqrt{1+\tan^2 \varphi}} = \frac{1}{\sqrt{1+L^2\omega^2/R^2}} = \frac{R}{\sqrt{R^2+L^2\omega^2}}. \tag{7.7.16}$$

From (7.7.13), we obtain for the current amplitude

$$I_0 = \frac{\mathcal{E}_0}{R\cos\varphi - L\omega \sin \varphi} = \frac{\mathcal{E}_0}{R(\cos \varphi + \tan \varphi \sin \varphi)}$$
$$= \frac{\mathcal{E}_0}{R} \cos \varphi = \frac{\mathcal{E}_0}{\sqrt{R^2+L^2\omega^2}}. \tag{7.7.17}$$

Here, we have successively used formulas (7.7.15) and (7.7.16). Thus, the problem formulated above is solved.

If the emf $\mathcal{E}_1$ is instantaneously switched off, current (7.7.12) does not immediately vanish but starts to at-

tenuate. Putting $\mathscr{E}_1 = 0$ in (7.7.10), we obtain the following equation for the transient process of attenuation:

$$L \frac{dI}{dt} + RI = 0. \tag{7.7.18}$$

This equation can be solved formally if we separate the variables. Namely, we transform (7.7.18) as follows:

$$\frac{dI}{I} = -\frac{R}{L} dt$$

and integrate both sides of the obtained equality:

$$\int \frac{1}{I} dI = -\frac{R}{L} \int dt + \text{const},$$

or

$$\ln I = -\left(\frac{R}{L}\right) t + \text{const}. \tag{7.7.19}$$

Let us now use the well-known formula

$$a = \ln e^a. \tag{7.7.20}$$

Then (7.7.19) will assume the form

$$\ln I = \ln e^{-(R/L)t} + \ln e^{\text{const}}$$
$$= \ln (e^{\text{const}} \cdot e^{-(R/L)t}) = \ln (\text{const}' \, e^{-(R/L)t}).$$

As a result, we obtain

$$I = \text{const}' \, e^{-(R/L)t}. \tag{7.7.21}$$

Over the time interval

$$\Delta t \gtrsim \frac{L}{R} \tag{7.7.22}$$

the current in the circuit vanishes almost completely.

Let us consider quite a different circuit. We speak of an **oscillatory circuit**, i.e. the circuit containing the inductance coil L and a capacitor of capacitance C. We shall neglect the resistance R of the circuit and supply a charge Q to the capacitor plates (Fig. 7.13). We aim at clarifying the phenomena occurring during the discharge of the capacitor. First of all, we must write the equation for the circuit

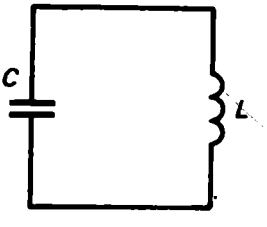


FIG. 7.13

under consideration. Using (7.5.8), (7.7.8), and Kirchhoff's law (7.3.10), we obtain

$$\frac{Q(t)}{C} = \mathscr{E}_{s.i}. \tag{7.7.23}$$

Bearing in mind that

$$I = \frac{dQ(t)}{dt} \tag{7.7.24}$$

(see Eq. (7.3.2)) and going over in the expression for $\mathscr{E}_{s.i}$ from the current $I(t)$ to the charges $Q(t)$, we can write Eq. (7.7.23) in the form

$$\frac{Q(t)}{C} = -\frac{1}{c^2} L \frac{d}{dt}\left(\frac{dQ(t)}{dt}\right) = -\frac{1}{c^2} L \frac{d^2Q(t)}{dt^2},$$

or

$$\frac{1}{c^2} L \frac{d^2Q(t)}{dt^2} + \frac{1}{C} Q(t) = 0. \tag{7.7.25}$$

Let us write this equation in the SI units. For this purpose, we shall use the conversion factors (see Appendix II, Table II.1):

$$Q \to \frac{1}{\sqrt{4\pi\varepsilon_0}} Q, \quad C \to \frac{1}{4\pi\varepsilon_0} C, \quad L \to \frac{4\pi}{\mu_0} L, \quad c^2 \to \frac{1}{\varepsilon_0\mu_0}.$$

As a result, we obtain

$$L \frac{d^2Q}{dt^2} + \frac{1}{C} Q = 0. \tag{7.7.25*}$$

This equation is of the type

$$\frac{d^2f}{dt^2} = -f. \tag{7.7.26}$$

The last equation describes a very important physical object, viz. a **harmonic oscillator**. In order to solve this equation, we must in fact find a function $f(t)$ which as a result of double differentiation transforms into itself with the opposite sign. Such a function is

$$f = \cos t. \tag{7.7.27}$$

Indeed, $df/dt = -\sin t$ and $d^2f/dt^2 = -\cos t = -f$.

Therefore, the solution of Eq. (7.7.25*) reduced to

$$\frac{d^2Q}{dt^2} = -\frac{1}{LC} Q \tag{7.7.28}$$

is a function of the form

$$Q(t) = \cos \omega t. \tag{7.7.29}$$

In order to find the unknown parameter ω, we substitute (7.7.29) into (7.7.28). For this purpose, we successively differentiate this expression twice:

$$\frac{dQ}{dt} = -\omega \sin \omega t,$$

$$\frac{d^2Q}{dt^2} = -\omega^2 \cos \omega t.$$

As a result, Eq. (7.7.28) becomes

$$-\omega^2 \cos \omega t = -\frac{1}{LC} \cos \omega t,$$

or

$$\omega^2 = \frac{1}{LC}. \tag{7.7.30}$$

The parameter ω is just the **natural frequency** of charge oscillations in the LC circuit. It is usually written in the form

$$\omega_0 = \frac{1}{\sqrt{LC}}. \tag{7.7.31}$$

Oscillations with ω_0 are undamped since we neglected the resistance R in which Joule heat losses could take place. However, for exciting oscillations in the circuit, an energy must be introduced into it (say, in the form of the energy of the electric field in the capacitor).

What will happen to natural oscillations in an LC circuit if we take into consideration its resistance R? The oscillations excited in the circuit will obviously soon attenuate. Let us derive the law of such a damped oscillation. First of all, the equation for the RLC circuit is

$$R \frac{dQ(t)}{dt} + \frac{1}{C} Q(t) = \mathscr{E}_{s.i}. \tag{7.7.32}$$

or

$$\frac{1}{c^2} L \frac{d^2Q(t)}{dt^2} + R \frac{dQ(t)}{dt} + \frac{1}{C} Q(t) = 0, \tag{7.7.33}$$

and has a similar form in the SI (the only difference being that the constant c^2 is missing),[2] i.e.

$$L \frac{d^2Q}{dt^2} + R \frac{dQ}{dt} + \frac{1}{C} Q = 0. \tag{7.7.33*}$$

Equation (7.7.32) is a generalization of Eq. (7.7.23).

It is natural to seek the solution of Eq. (7.7.33*) in the form

$$Q(t) = Ae^{-\gamma t} \cos \omega t, \tag{7.7.34}$$

where A, γ, and ω are unknown constants.

Let us find the first and the second derivative of function (7.7.34):

$$\frac{dQ}{dt} = A(-\gamma e^{-\gamma t} \cos \omega t - \omega e^{-\gamma t} \sin \omega t)$$

$$= -Ae^{-\gamma t} (\gamma \cos \omega t + \omega \sin \omega t),$$

$$\frac{d^2Q}{dt^2} = Ae^{-\gamma t} [(\gamma^2 - \omega^2) \cos \omega t + 2\gamma\omega \sin \omega t].$$

Substituting $Q(t)$, dQ/dt, and d^2Q/dt^2 into (7.7.33*) and cancelling out the factor $Ae^{-\gamma t}$, we obtain

$$L[(\gamma^2 - \omega^2) \cos \omega t + 2\gamma\omega \sin \omega t]$$

$$- R(\gamma \cos \omega t + \omega \sin \omega t) + \frac{1}{C} \cos \omega t = 0.$$

This expression is valid for any moment of time t if the coefficients of $\cos \omega t$ and $\sin \omega t$ are zero. As a result, we obtain two conditions

$$L(\gamma^2 - \omega^2) - R\gamma + \frac{1}{C} = 0 \tag{7.7.35}$$

and

$$L \cdot 2\gamma - R = 0. \tag{7.7.36}$$

From (7.7.36), we obtain

$$\gamma = \frac{R}{2L}. \tag{7.7.37}$$

[2] The reader can easily verify this by using the appropriate conversion factors.

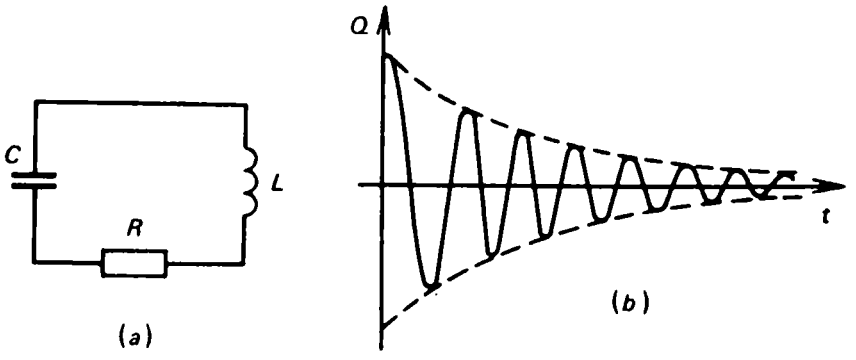


FIG. 7.14

Substituting the expression for γ into (7.7.35), we determin

$$\omega^2 = \frac{1}{LC} - \frac{R^2}{4L^2}. \qquad (7.7.38)$$

In the limit $R \to 0$, Eq. (7.7.38) is transformed to (7.7.30). The quantity γ is known as the **damping decrement**. A comparison of the damping decrement (7.7.37) with the similar decrement appearing in (7.7.21) indicates that the damping itself is only due to the resistance R and the "inertia" of coil L. The capacitance C has no effect on damping. The nature of damping in the RLC circuit (Fig. 7.14*a*) is shown in Fig. 7.14*b*.

In order to make oscillations in an RLC circuit undamped, the energy losses in it must be compensated from an external source. The introduction of an external periodic emf $\mathscr{E}_1$ into the RLC circuit will produce in the oscillatory circuit undamped oscillations with a frequency equal to the frequency of variation of the emf. In this case, the equation for oscillations of the charge $Q(t)$ differs from (7.7.33*) in the presence of the emf on the right-hand side:

$$L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = \mathscr{E}_0 \cos \omega t. \qquad (7.7.39)$$

Let us go over from the equation for charge to the equation for current. Before doing that, we obtain from (7.7.24)

$$dQ = I\,dt,$$

or

$$Q = \int I\,dt. \qquad (7.7.40)$$

As a result, Eq. (7.7.39) becomes

$$L\frac{dI}{dt}+RI+\frac{1}{C}\int I\,dt=\mathscr{E}_0\cos\omega t. \tag{7.7.41}$$

Considering only steady-state conditions of forced oscillations, we shall seek the solution of Eq. (7.7.41) in the form

$$I(t)=I_0\cos(\omega t+\varphi), \tag{7.7.42}$$

where the constants I_0 and φ are to be determined.

Let us evaluate the integral appearing in (7.7.41):

$$\int I\,dt=\int I_0\cos(\omega t+\varphi)\,dt$$

$$=\frac{I_0}{\omega}\int\cos(\omega t+\varphi)\,d(\omega t+\varphi)=\frac{I_0}{\omega}\sin(\omega t+\varphi).$$

As a result, Eq. (7.7.41) with function (7.7.42) will be written as

$$-LI_0\omega\sin(\omega t+\varphi)+RI_0\cos(\omega t+\varphi)$$

$$+\frac{I_0}{\omega C}\sin(\omega t+\varphi)-\mathscr{E}_0\cos\omega t=0.$$

Let us expand the expressions for sine and cosine of the sum of two angles. Equating to zero the coefficients of $\sin\omega t$ and $\cos\omega t$, we obtain two conditions:

$$\left(\frac{1}{\omega C}-L\omega\right)\cos\varphi-R\sin\varphi=0 \tag{7.7.43}$$

and

$$\left(\frac{I_0}{\omega C}-LI_0\omega\right)\sin\varphi+RI_0\cos\varphi-\mathscr{E}_0=0. \tag{7.7.44}$$

Using (7.7.43), we determine φ:

$$\tan\varphi=\left(\frac{1}{\omega C}-\omega L\right)\frac{1}{R}. \tag{7.7.45}$$

The expression for I_0 can be obtained from (7.7.44), but before that we write the expressions for

$$\sin\varphi=\frac{\tan\varphi}{\sqrt{1+\tan^2\varphi}}=\frac{1/(\omega C)-L\omega}{\sqrt{R^2+[1/(\omega C)-\omega L]^2}}$$

and

$$\cos\varphi = \frac{1}{\sqrt{1+\tan^2\varphi}} = \frac{R}{\sqrt{R^2+[1/(\omega C)-\omega L]^2}}.$$

Substituting these expressions into (7.7.44), we obtain

$$I_0 = \frac{\mathscr{E}_0}{\sqrt{R^2+[1/(\omega C)-\omega L]^2}}. \qquad (7.7.46)$$

The current amplitude I_0 attains the maximum value when the radicand in (7.7.46) is at the minimum, i.e. when

$$\left(\frac{1}{\omega C}-\omega L\right)=0.$$

Hence

$$\omega = \frac{1}{\sqrt{LC}} = \omega_0. \qquad (7.7.47)$$

When the frequency ω of external perturbation coincides with the natural frequency ω_0 of oscillations in the circuit (see Eq. (7.7.31)), the amplitude of current oscillations becomes $\mathscr{E}_0/R$. As $R \to 0$, the current amplitude $I_0 \to \infty$. This phenomenon is known as **resonance.**

7.7.4. Generation of Electromagnetic Waves. The presence of a high-frequency current in an *RLC* circuit indicates that conduction electrons move with a high acceleration (acceleration is proportional to the squared frequency, see Sec.3.8.3). But any accelerated motion of charges is accompanied by the emission of electromagnetic waves, the radiation intensity being proportional to the fourth power of the frequency (see Eq. (3.8.7)). The natural frequencies of an *LC* circuit increase with decreasing capacitance and inductance (see Eq. (7.7.31)). This suggests a method for generating high-frequency electromagnetic waves: *we must operate with circuits having both small inductance and capacitance.*

Let us derive an expression for the capacitance C of a capacitor used in an oscillatory circuit (Fig. 7.15*a*). We proceed from relation (7.5.8)

$$C = \frac{Q}{\Delta\varphi}.$$

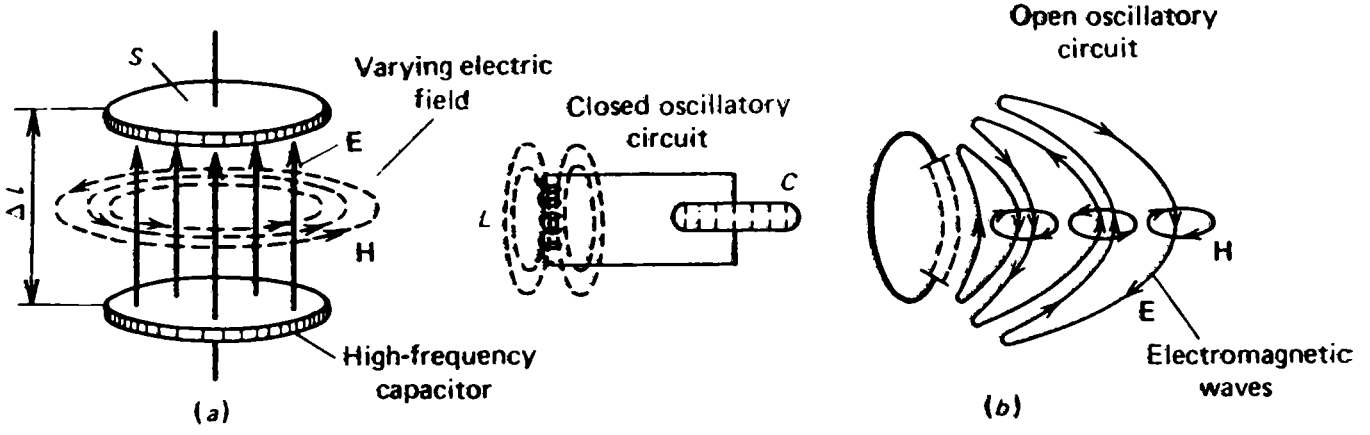

S
Varying electric field
E
H
Δl
High-frequency capacitor
(a)
Closed oscillatory circuit
L
C
Open oscillatory circuit
H
E
Electromagnetic waves
(b)

FIG. 7.15

The potential difference $\Delta\varphi$ can be determined from (3.4.14):

$$\Delta\varphi = E\,\Delta l,$$

where Δl is the separation between the capacitor plates. The electric field strength E in the capacitor has already been calculated by us (see Eq. (3.4.5)):

$$E = 4\pi\sigma.$$

Here, σ is the surface charge density which determines the charge in the initial formula:

$$Q = \sigma S,$$

where S is the area of the capacitor plates. Combining these expressions, we obtain

$$C = \frac{Q}{\Delta\varphi} = \frac{\sigma S}{E\,\Delta l} = \frac{\sigma S}{4\pi\sigma\,\Delta l} = \frac{S}{4\pi\,\Delta l}. \qquad (7.7.48)$$

This formula shows that in order to reduce the capacitance of a capacitor, we must reduce the area of the plates and increase the distance between them. As regards the inductance L, a transition to small values of L means that coils should be replaced by linear conductors. As a result, the closed low-frequency oscillatory circuit is transformed into an open high-frequency oscillatory circuit (Fig. 7.15*b*).

7.7.5. Some Laws of Optics and the Velocity of Propagation of Electromagnetic Waves in a Medium. Reflection and Refraction of Waves. In this textbook, optical phenomena are considered in many different sections. This is done because, firstly, light is inseparably linked with many physical phenomena, and secondly, light is a multifaceted object whose properties become clear as we go over to deeper levels of analysis (from classical physics to relativistic physics and from relativism to quantization).

What do we know about light? First of all, it should be recalled that (see Sec. 3.9) light is the form of electromagnetic wavès corresponding to a certain frequency range. In Sec. 3.9, we considered the problems connected with the laws of light propagation in vacuum, in particular, interference and diffraction. The conditions under which a transition can be made from the wave pattern of light propagation to the ideas of ray (geometrical) optics were specified.

Considering the sources of radiation (light), we analyzed in Sec. 3.7 the classical betatron emission of free electrons in a magnetic field in the visible frequency range. The mechanism of glow of gaseous nebulas in interstellar space and, in particular, the glow of the remnants of supernovas, viz. their envelopes, during flashes is just of this type. In Sec. 4.7, we considered the discrete emission of bound electrons in atoms, emerging during quantum transitions in excited atoms. We know that it were these line emission spectra of atoms that formed the basis of the spectral analysis of substances not only in the terrestrial conditions but also in far and near cosmic space.

In Sec. 4.7, we considered the conditions under which the photon pattern of a radiation field is clearly manifested, indicated the reason behind the possible limiting transition from the photon pattern of the radiation field to the classical electromagnetic waves, and analyzed the experimental foundations for the quantization of the radiation field, i.e. the photoelectric effect and the Compton effect.

Section 5.9 was devoted to thermal radiation of macroscopic bodies. There we used the idea about thermal radiation as an ideal photon gas. It was found that this equilibrium photon gas is described by Planck's formula which includes temperature as a parameter. Analyzing the frequency distribution of energy in solar and stellar radiations and assuming this radiation to be of thermal origin, it is possible to determine the temperatures of these remote objects. Moreover, the discovery that the isotropic electromagnetic relict radiation (see Sec. 9.8.1) has a Planck-type spectrum and is of thermal origin has made it possible for astrophysicists and cosmologists to look into the past history of our Universe and learn some facts about the state of matter at that time (for example, about the temperature). During the expansion of the Universe, the effective temperature of this "residual" thermal radiation has decreased to 3 K.

Finally, in Sec. 7.4, we considered such optical phenomena as light absorption in semiconducting crystals and the formation of exciton states. In this section, we also studied the intrinsic photoelectric effect in semiconductors, as well as the principle of operation of solar cells.

Concluding this review of optical phenomena, we note the exceptionally important role of optical phenomena in the

evolution of physics and, in particular, such fundamental physical theories as the theory of relativity and the quantum theory. As regards the relation between optics and modern engineering, we should mention such its trends as lasers, laser communication, holography, optoelectronics, fibre optics, and optical computers of future.

Continuing the analysis of optical phenomena, we shall speak, for the sake of generality, about electromagnetic waves since light is a special case of such waves. We are now facing the problem of the effect of the medium on the course of optical phenomena.

It has already been noted in Sec. 7.7.1 that external varying electromagnetic fields may considerably affect the properties of a substance. This effect consists, for example, in the strong dependence of the polarization of a substance, associated with orientational oscillations of dipole and magnetic moments of individual particles, on the frequency of the external field.

There is a reverse effect of a substance (medium) on the properties of propagation of varying electromagnetic fields in them. For example, electromagnetic waves undergo reflection and refraction at the boundary between two media. The velocity of propagation of electromagnetic waves in a substance differs from that in vacuum (it is lower than the limiting value).

Unlike in vacuum, various **nonlinear electrodynamic effects** can be observed in a medium. As a matter of fact, the basic equations of electrodynamics are linear (see Sec. 3.3) and obey the principle of superposition of fields. According to this principle, electromagnetic waves are not scattered by one another. On the other hand, electromagnetic waves of very high intensity of the fields (like laser beams) affect the medium so strongly that its permittivity and permeability become functions of the fields. In turn, these variations of $\varepsilon(\omega, E)$ and $\mu(\omega, H)$ affect the electromagnetic waves: the waves of new frequencies are generated, as a result of "interaction of the initial waves".

Another example of the principal role of the medium in electrodynamic phenomena is that charges moving with a constant velocity in vacuum do not emit (see Sec. 3.7.2), while the field of a high-energy electron flying into a transparent insulator with a velocity exceeding the speed of

light in this medium causes its emission. The leading wave front of this radiation resembles the shock wave emerging in a medium when bodies move in it with supersonic velocities (see Fig. 6.17*b*). This radiation is known as the **Cherenkov-Vavilov radiation.**

Let us now analyze quantitatively the behaviour of electromagnetic waves (in particular, light) in a medium. Why does light change the velocity of propagation as it passes from vacuum to a transparent insulating medium? As a matter of fact, in addition to the polarization of molecules (or atoms) of the medium, the rapidly varying electric field of a light wave causes equally rapid oscillations of induced dipole moment. These forced oscillations of all the dipoles of the medium are matched since the front of the light wave embraces sufficiently large regions at the same time. And as the vibrational motion of the dipole charges is accelerated, a radiation field emerges in the medium. This radiation field of the medium itself interferes with the original field of the light wave. As a result of interference, a phase shift emerges in the light wave. On the whole, this is manifested as a change in the phase velocity of light in the medium (see Sec. 6.8.2 on the phase velocity and group velocity of wave processes).

It is well known that the electric polarizability of a substance determines the value of the permittivity of the medium (see Sec. 7.5.3). Therefore, it is natural to expect that the phase velocity v_{phase} of light in a homogeneous insulating medium is expressed in terms of the speed of light c in vacuum, the permittivity ε, and, in the general case, the permeability μ, namely,

$$v_{\text{phase}} = \frac{c}{\sqrt{\varepsilon\mu}}. \tag{7.7.49}$$

In the case of vacuum, $\varepsilon = \mu = 1$, and the wave velocity is c. In insulators, $\varepsilon > 1$, i.e. the wave velocity in them is smaller than c.

Since the velocity of electromagnetic waves changes at the boundary between vacuum and an insulating medium, this, naturally, leads to a change in the direction of wave propagation, i.e. to its **refraction** at the interface (Fig. 7.16). The ratio of the velocity c of propagation of electromagnetic

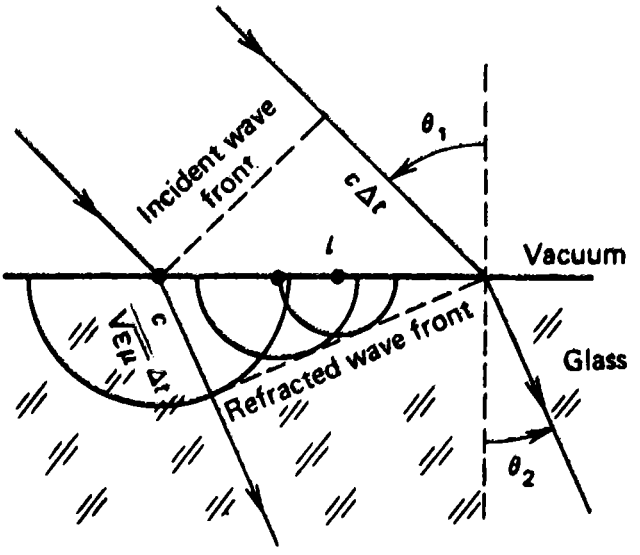


FIG. 7.16

waves in vacuum to their velocity $c/\sqrt{\varepsilon\mu}$ in the medium is called the **refractive index** of the medium:

$$n = \sqrt{\varepsilon\mu}. \tag{7.7.50}$$

Let us find the relation between the angle of incidence θ_1 and the angle of refraction θ_2 of a light wave (see Fig. 7.16). For this purpose, we consider two right triangles formed by the fronts of the incident and refracted waves and the common hypotenuse l coinciding with the vacuum-medium interface. We have

$$c\,\Delta t = l \sin\theta_1$$

and

$$\frac{c}{\sqrt{\varepsilon\mu}}\,\Delta t = l \sin\theta_2.$$

Dividing the first equation by the second and taking into account (7.7.50), we obtain

$$n = \frac{\sin\theta_1}{\sin\theta_2}. \tag{7.7.51}$$

This is **Snell's law of refraction.**

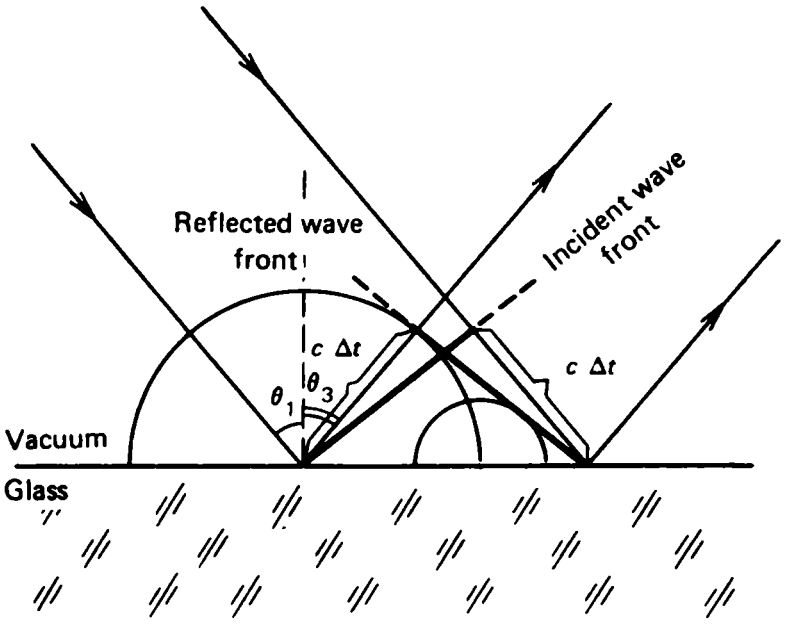


FIG. 7.17

Experiments show that at the vacuum-medium interface, as well as at the interface between two transparent insulating media, an incident wave simultaneously gives rise to a refracted and a reflected wave. The formation of the reflected wave is illustrated in Fig. 7.17. Figures 7.16 and 7.17 should be combined into one and are considered separately only for the sake of convenience. Using the same arguments as for the refracted wave, we obtain

$$c\,\Delta t = l \sin \theta_1,\ c\,\Delta t = l \sin \theta_3,$$

or

$$\sin \theta_1 = \sin \theta_3.$$

This equality is valid when

$$\theta_1 = \theta_3. \tag{7.7.52}$$

For an incident and a reflected wave, the angle of incidence θ_1 *is equal to the angle of reflection* θ_3.

It should be noted that relations (7.7.51) and (7.7.52) automatically follow from the analysis of the boundary conditions for the fields since Maxwell's equations must be satisfied at the interface.

7.7.6. Refractive Index of Insulators. Dispersion and Absorption of Light. Ordinary insulators may exhibit dia- and paramagnetic properties, and their permeability μ is very close to unity. The values $\mu \gg 1$ are observed for insulators with ferromagnetic structure (ferrites), but for optical frequencies of electromagnetic waves, their permeability $\mu = 1$. This circumstance is due to the fact that the frequency dependence of μ for ferromagnetic insulators is manifested quite soon (owing to the inertia of large domains) and is accompanied by a drop in the value of μ almost to unity. Thus, formula (7.7.50) can be written in a very good approximation as follows:

$$n^2 = \varepsilon. \tag{7.7.53}$$

Using expression (7.5.18) for ε, we can write (7.7.53) in the following form:

$$n^2 = 1 + \alpha N, \tag{7.7.54}$$

where N is the number of atoms per unit volume, and α is the atomic polarizability. The latter quantity is connected with the induced dipole moment of the atom through the formula

$$\mathbf{d} = \frac{1}{4\pi} \alpha \mathbf{E}. \tag{7.7.55}$$

Here, we have used relations (7.5.13) and (7.5.17). In turn,

$$\mathbf{d} = e\mathbf{r} \tag{7.7.56}$$

see Sec. 3.4.4), where e is the charge of an atomic electron, and $\mathbf{r}$ is its displacement in the electric field of the wave.

Let us consider the model of an atom in which an electron performs forced oscillations under the action of the periodic field of a light wave. We write the equation of motion for such coupled oscillations:

$$m \frac{d^2\mathbf{r}}{dt^2} = -m\omega_0^2 \mathbf{r} + e\mathbf{E}. \tag{7.7.57}$$

Here, the right-hand side contains the "elastic" restoring force (see Sec. 7.5.3) and the force perturbation produced by the electric field of the light wave. Let the wave be linearly

polarized, and the electron oscillate along the field. Then Eq. (7.7.57) will be written in the form

$$\frac{d^2r}{dt^2} + \omega_0^2 r = \frac{e}{m} E_0 \cos \omega t. \tag{7.7.58}$$

We devoted much attention to the solution of this type of equations in Sec. 7.7.3. It is natural to seek the displacement r of the electron in the form

$$r = r_0 \cos \omega t. \tag{7.7.59}$$

The electron oscillates synchronously with the field. Substituting (7.7.59) into (7.7.58), we obtain

$$r_0 = \frac{e/m}{\omega_0^2 - \omega^2} E_0,$$

or

$$\mathbf{r} = \frac{e/m}{\omega_0^2 - \omega^2} \mathbf{E}. \tag{7.7.60}$$

Let us now retrace the chain of formulas from (7.7.56) to (7.7.54). For this purpose, we substitute (7.7.60) into (7.7.56):

$$\mathbf{d} = \frac{e^2/m}{\omega_0^2 - \omega^2} \mathbf{E}. \tag{7.7.61}$$

Comparing this formula with (7.7.55), we obtain

$$\alpha = 4\pi \frac{e^2/m}{\omega_0^2 - \omega^2}. \tag{7.7.62}$$

Finally, substituting (7.7.62) into (7.7.54), we can write

$$n^2 = 1 + 4\pi N \frac{e^2/m}{\omega_0^2 - \omega^2}. \tag{7.7.63}$$

It should be noted that the analysis carried out in this section is of qualitative nature. Therefore, it is not surprising that all the formulas containing the factor $1/(\omega_0^2 - \omega^2)$ become meaningless when $\omega \to \omega_0$. These formulas may become correct if we take into account the damping of oscillations of the atomic electron under the effect of its interaction with the neighbouring atoms. Dependence (7.7.63) is graphically represented in Fig. 7.18. The real curve $n(\omega)$ taking the damping into account is shown in Fig. 7.19.

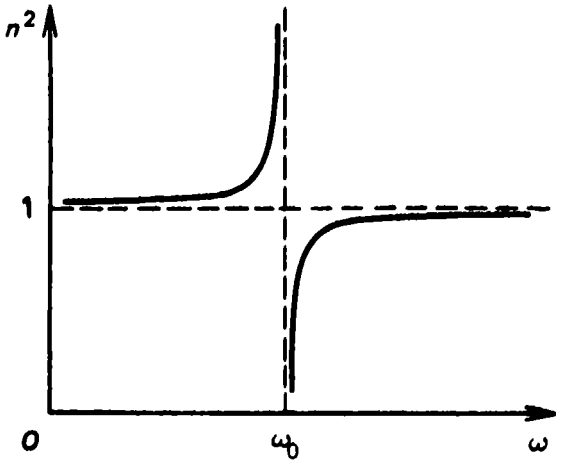


FIG. 7.18

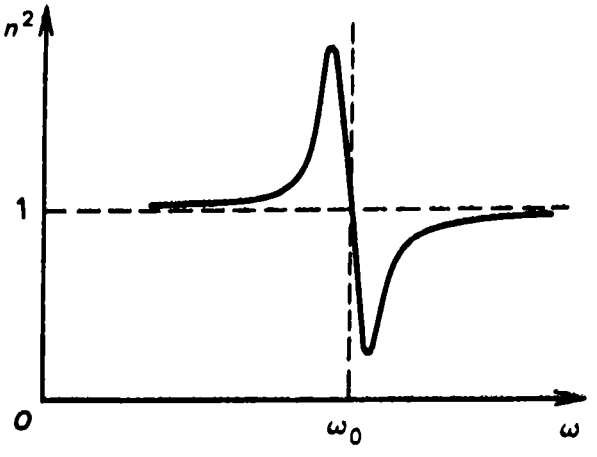


FIG. 7.19

Since, according to Eq. (7.7.63), the refractive index $n(\omega)$ is a function of frequency, the natural white light, which is a set of waves with different close frequencies, undergoes decomposition (dispersion) into its "coloured" components during its propagation through an insulator. It should be borne in mind that the "eigenfrequencies" ω_0 of atomic electrons in (7.7.63) are rather high (for example, for hydrogen in the ground state they belong to the ultraviolet region of the spectrum, see Fig. 4.15). For this reason, in particular, blue rays which correspond to a higher frequency in comparison with red rays, have, according to (7.7.63), a larger refractive index.

When the frequencies ω of electromagnetic radiation are close to the "eigenfrequencies" ω_0 of a substance, a resonance absorption of the radiation by the substance takes place. For this reason, for example, ordinary glass does not transmit ultraviolet radiation.

7.7.7. Refractive Index of Metals. Skin Effect. Transparency of Metals to Hard Radiation. Since metals, unlike insulators, contain free electrons, these electrons absorb practically the entire energy of an external varying electric field. Putting $\omega_0 = 0$ in (7.7.63) (there is no "elastic" restoring force in the equation of motion (7.7.57) for free electrons), we obtain

$$n^2 = 1 - 4\pi \frac{N_e e^2}{m\omega^2} , \tag{7.7.64}$$

or

$$n^2 = 1 - \frac{\omega_p^2}{\omega^2} , \tag{7.7.65}$$

where

$$\omega_p = \sqrt{\frac{4\pi N_e e^2}{m}} \tag{7.7.66}$$

is the plasma frequency for free electrons (see Sec. 8.7), and N_e is the number of free electrons per unit volume.

Let us consider two limiting cases. For

$$\omega^2 < \omega_p^2, \tag{7.7.67}$$

the value of n^2 becomes negative, i.e. the refractive index becomes imaginary. This means that an electromagnetic wave does not pass through a substance and is abruptly quenched (absorbed) in a thin surface layer in the frequency range under consideration.

What is the mechanism of energy absorption of a wave in a metal? If we compare (7.7.66) with (7.2.6), condition (7.7.67) becomes

$$\omega^2 < \frac{4\pi\sigma}{\tau} , \tag{7.7.68}$$

where σ is the electron conductivity, and τ is the mean time between collisions of free electrons with phonons (lattice

vibrations). Thus, when the frequency of the field is such that

$$\omega \ll \frac{1}{\tau}, \tag{7.7.69}$$

a conduction electron has time to undergo multiple collisions with the lattice during a wave period. As a result, the entire energy of the field absorbed by the electrons will ultimately be converted into heat. In actual practice, this is manifested in **the skin effect**: the surface layer of a metal sample in a high-frequency field is strongly heated, while the core of the sample remains cold.

In the opposite limiting case, i.e. when the frequencies of electromagnetic waves are such that

$$\omega \gg \omega_p, \tag{7.7.70}$$

the refractive index (7.7.65) becomes real-valued, and hence the metal becomes transparent. What is the order of magnitude of the frequency ω_p of oscillations of the number density of free electrons? Since $N_e \sim 10^{28}$ electrons/m^3 in metals (see Sec. 5.10.6), by substituting the values of charge e and mass m of the electron into (7.7.66), we obtain

$$\omega_p \sim 3 \times 10^{15} \text{ s}^{-1}$$

(it should be recalled that we must substitute the quantities expressed in the CGS units into (7.7.66) and then convert the result into the SI units). This frequency corresponds to the ultraviolet radiation. Taking into account (7.7.70), we may conclude that metals are transparent to X-rays. This is confirmed by experiments.

The reason behind the transparency of metals to hard short-wave radiation lies in that inequality (7.7.69) in this case is replaced by the reverse inequality. This means that the oscillations of conduction electrons in an external rapidly varying field occur practically without collisions with "neighbours".

7.7.8. Nonlinear Optics Effects. If the electric field strengths in a light wave are considerably smaller than intra-atomic field strengths, the effect of light on the optical properties of the medium can be neglected. By way of an example, let

us estimate the field strength in a hydrogen atom. For this purpose, we shall use expression (3.4.1) which gives

$$E = \frac{e}{r^2}, \quad E = 6 \times 10^{11} \text{ V/m}$$

(we substitute the values expressed in the CGS units into this formula and then convert the result into the SI units). Such an electric field acts on the electron in a hydrogen atom.

Light fields from ordinary thermal sources (like incandescent lamps) are extremely weak in comparison with values of the order of 10^{11} V/m, and we shall call such fields weak. Lasers make it possible to obtain light waves with the field strength of 10^8-10^{10} V/m. These fields are comparable with internal fields in a medium, and we shall call such fields strong.

Weak light fields cannot noticeably affect atomic electrons or the optical characteristics of a medium. For example, the refractive index will remain unchanged, i.e. will not depend on wave intensity.

If such characteristics of a medium as ε, μ, and n are independent of the light field strength, the light wave propagating in the medium should be described by a system of linear differential equations. For this reason, the laws governing the optical phenomena in the medium and obtained on the basis of the equations mentioned above are known as **linear.**

In order to demonstrate the nature of the effect of high-intensity light on the optical properties of a medium, we shall consider again (as in Sec. 7.7.6) the model of an atom in which an electron performs forced oscillations under the effect of the periodic field of a light wave. Since the field is strong, the oscillations cannot be treated as harmonic any longer, and for the "elastic" restoring force we must take the following nonlinear expression:

$$\mathbf{F}(\mathbf{r}) = -m\omega_0^2\mathbf{r} - A\mathbf{r}^3. \tag{7.7.71}$$

In Sec. 7.7.6, we limited ourselves only to the linear dependence $\mathbf{F}(\mathbf{r})$, i.e. only the first linear term was present on the right-hand side of Eq. (7.7.71). In Sec. 5.10.3, it was explained why the cubic term should be taken as an anharmonic (nonlinear) term (see formula (5.10.7)).

We shall consider the frequency range ω of the light wave field

$$\mathbf{E} = \mathbf{E}_0 \cos \omega t \tag{7.7.72}$$

away from the resonance frequencies ω_0 of the atom. Then we can neglect the damping of oscillations of the atomic electron and write its equation of motion in the form

$$m \frac{d^2\mathbf{r}}{dt^2} = e\mathbf{E}(t) - m\omega_0^2 \mathbf{r} - A\mathbf{r}^3 \tag{7.7.73}$$

(cf. Eq. (7.7.57)).

Let the light wave field be strong but still considerably weaker than the internal field. In this case, the nonlinear term in (7.7.73) can be treated as small, and the equation can be solved by the **method of successive approximations.** The essence of this method lies in that the solution is sought in the form

$$\mathbf{r} = \mathbf{r}_{(1)} + \mathbf{r}_{(2)}, \tag{7.7.74}$$

where $\mathbf{r}_{(2)} \ll \mathbf{r}_{(1)}$.

The function $\mathbf{r}_{(1)}$ (the **first approximation**) satisfies the equation

$$m \frac{d^2\mathbf{r}_{(1)}}{dt^2} + m\omega_0^2 \mathbf{r}_{(1)} = e\mathbf{E}(t).$$

This equation coincides with Eq. (7.7.57) whose solution is known to us (see Eq. (7.7.60)):

$$\mathbf{r}_{(1)} = \frac{e/m}{\omega_0^2 - \omega^2} \mathbf{E}. \tag{7.7.75}$$

The **second approximation** $\mathbf{r}_{(2)}$ is sought from the equation

$$m \frac{d^2\mathbf{r}_{(2)}}{dt^2} + m\omega_0^2 \mathbf{r}_{(2)} = -A\mathbf{r}_{(1)}^3.$$

Substituting expression (7.7.75) for $\mathbf{r}_{(1)}$ into this equation, we obtain

$$m \frac{d^2\mathbf{r}_{(2)}}{dt^2} + m\omega_0^2 \mathbf{r}_{(2)} = -\frac{A}{(4\pi)^3} \alpha^3(\omega) \mathbf{E}^3(t), \tag{7.7.76}$$

where $\alpha(\omega)$ is the atomic polarizability (see Eq. (7.7.62)). Using in (7.7.76) the explicit form of the light wave field

(7.7.72) and the trigonometric identity

$$\cos^3 \omega t = \frac{1}{4}(3\cos \omega t + \cos 3\omega t),$$

we obtain

$$\frac{d^2 \mathbf{r}_{(2)}}{dt^2} + \omega_0^2 \mathbf{r}_{(2)}$$

$$= -\frac{A}{4(4\pi)^3 m} \alpha^3(\omega) \mathbf{E}_0^3 (3\cos \omega t + \cos 3\omega t). \qquad (7.7.77)$$

As a result, the second approximation $\mathbf{r}_{(2)}$, as well as Eq. (7.7.77), consists of two terms one of which oscillates with the same frequency ω as the light wave field $\mathbf{E}(t)$, and the other term oscillates with a frequency thrice as high (3ω).

Thus, the initial equation (7.7.73) for the displacement $\mathbf{r}(t)$ of an anharmonic oscillator has the solution

$$\mathbf{r}(t) = \mathbf{r}_{(1)}^{(\omega)}(t) + \mathbf{r}_{(2)}^{(\omega)}(t) + \mathbf{r}_{(2)}^{(3\omega)}(t),$$

which in explicit form can be written as

$$\mathbf{r}(t) = \frac{\alpha(\omega, E_0)}{e} \mathbf{E}_0 \cos \omega t + \frac{\alpha(3\omega, E_0)}{e} \mathbf{E}_0 \cos 3\omega t,$$

where

$$\alpha(\omega, E_0) = \alpha(\omega) - \frac{3A\alpha^3(\omega) E_0^2}{4me^2(\omega_0^2 - \omega^2)}$$

and

$$\alpha(3\omega, E_0) = -\frac{A\alpha^3(\omega) E_0^2}{4me^2(\omega_0^2 - 9\omega^2)}.$$

What is the meaning of the obtained results? Atomic electrons oscillating under the influence of the strong light wave of frequency ω re-emit electromagnetic waves (see Sec. 3.8.3) with the same frequency ω and an additional frequency 3ω. Further, since the atomic polarizability α is connected with the refractive index n (see Eq. (7.7.54)), $n(\omega, E_0)$ is a function of frequency and field (to be more precise, a function of the squared frequency and field strength). The refractive index, in turn, is associated with the permittivity ε (see Eq. (7.7.53)), i.e. $\varepsilon(\omega, E_0)$ is also a function of frequency and field strength of the light wave.

If a light beam of frequency ω, emitted by an ordinary thermal emitter, is incident on a transparent insulating crystal, the light passing through the crystal will have the same frequency ω. If a laser beam of frequency ω passes through this crystal, in the emerging beam there will be two components with frequencies ω and 3ω.[3]

Among other nonlinear optics effects, we shall mention the variation of the angle of refraction of a light beam at the interface between two media with the beam intensity, the compression of light beams during propagation in a medium **(self-focussing of light)**, and reversible clearing up (or, on the contrary, blackout) of a medium under the action of an incident light beam.

A systematic analysis of nonlinear optics effects has become possible only after the development of quantum optical generators, viz. **lasers.**

7.7.9. Lasers. In thermal sources of light, the microscopic mechanism of emission is reduced to spontaneous emission (see Sec. 4.7.2) of excited atoms of the medium. In spontaneous transitions, individual atoms emit independently of one another. As a result, the emitted radiation is characterized by a wide spread in frequency, direction of propagation, and polarization, i.e. it is an **incoherent radiation.** In view of incoherence of the waves emitted during spontaneous transitions, the mean intensity of the waves emitted by all the atoms is equal to the sum of the intensities of the waves emitted by individual atoms. In such light sources, the electric field strength $E \sim 0.1\text{-}1.0$ kV/m.

A **coherent radiation** must consist of identical waves. This means that their directions of propagation coincide, the orientations of electric and magnetic field vectors are identical, and the frequency is the same. Moreover, the phases of waves emitted by all the atoms must be matched (in particular, coincide).

Such a coherent emission by the atoms is characterized by an increase in the resultant amplitude of their radiation in proportion to the number of atoms, and the radiation intensity turns out to be proportional to the squared number of atoms.

[3] In a real experiment, special crystals are required, as well as a definite polarization of the beam, crystal orientation, and so on.

In order to obtain high-intensity coherent light beams, induced emission should be used. It should be recalled that the origin of the induced emission of atoms (see Sec. 4.7.2) is as follows. If a photon of the same type as that emitted by an atom is incident on the excited atom, the probability of emission of an identical photon by the atom sharply increases. Since photons are Bose particles, they can be accumulated in the same quantum state in an unlimited amount.

The attempts to employ the induced emission of atoms in developing the sources of high-intensity coherent light were foiled due to the following complication. In ordinary equilibrium states of a substance, the number of atoms of the medium "filling" different energy levels sharply drops as we go over to higher energy levels (see Fig. 5.7). Even if the temperature of the medium increases unlimitedly, the particle distribution among energy levels just becomes uniform (see Fig. 5.8).

An active medium with an inverse population of energy levels of the system is required for this purpose (see Fig. 5.8 corresponding to the case of negative temperatures). Such states can be created, for example, by an additional radiation **(pumping radiation)**.

Let us consider a three-level quantum system of a ruby crystal with a chromium admixture. The three energy levels of chromium atoms are used for creating a state with a negative temperature. The operation principle of such a laser is illustrated in Fig. 7.20.

In order to transfer chromium atoms from an unexcited state E_0 to an excited state E_2, gas-discharge lamps operat-

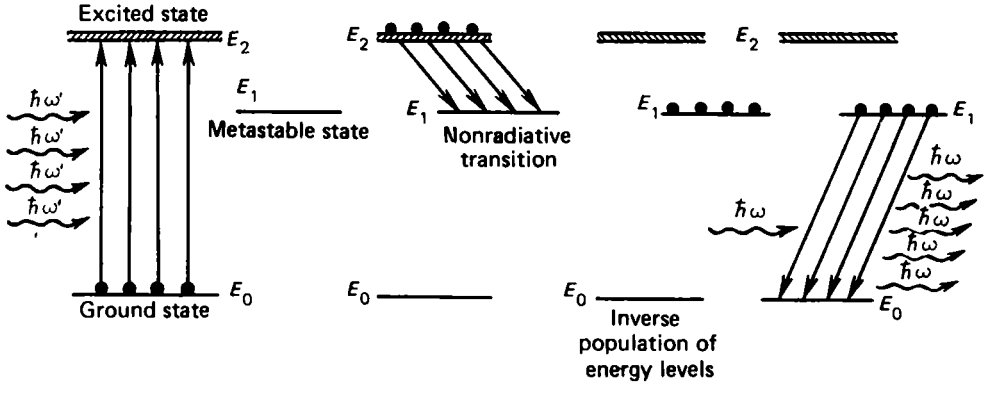


FIG. 7.20

ing in a pulsed mode are used for pumping. A gas discharge is initiated with the help of a special appliance containing capacitors of capacitance up to 1 mF. The capacitors are charged from a d.c. source of several thousand volts and then discharged through a gas-discharge tube. The pulse duration is of the order of 10^{-3} s, while the energy consumed during a pulse is of the order of 10^3 J. The energy of the discharge is transformed into the energy of pumping radiation.

The pumping radiation lies in the blue and green regions of the spectrum. The energy of pumping radiation is spent to excite chromium atoms which go over from the state E_0 to the state E_2.

As regards the state E_2, it should be noted that since its lifetime is short ($\sim 10^{-8}$ s), it corresponds to the whole "band" of states, i.e. the energy level has a "finite width". There exists a general quantum-mechanical relation between the broadening ΔE of the energy level and the lifetime Δt of this state: $\Delta E\, \Delta t \sim \hbar$. This circumstance is very important for the operation of a laser since the optical spectrum of the pumping lamp has a considerable broadening, and the broadening of the excited states E_2 is very advantageous for increasing the efficiency of the lamp.

During a time of the order of 10^{-8} s, chromium atoms perform a nonradiative transition to a metastable (long-lived) state E_1, having supplied the excess of their energy to crystal lattice vibrations.

The existence of the long-lived ($\sim 10^{-3}$ s) metastable state E_1 makes it possible to "accumulate" chromium atoms in this state and thus to create a higher population of the metastable level E_1 as compared with the ground level E_0. For creating such a population inversion, more than half the chromium atoms are transferred to the level E_1 over the time of the order of 10^{-3} s.

The presence of a **resonator**, viz. two plane-parallel mirrors at the end faces of the ruby crystal, makes it possible to sharply increase the number of photons of induced radiation propagating along the resonator axis.

The photon shower of induced radiation starts when a chromium atom from the metastable state spontaneously emits a photon parallel to the resonator axis since photons emitted in other directions rapidly leave the crystal. The

multiple reflection from the mirrors leads to an avalanche growth in the number of photons in the induced radiation from chromium atoms.

One of the mirrors of a laser crystal is made semitransparent, which ensures the emergence of a high-intensity coherent laser beam from the crystal. The radiation from a ruby laser corresponds to the red region of the spectrum, the degree of monochromaticity being extraordinarily high: $\Delta\lambda/\lambda \sim 10^{-4}$ (it should be noted for comparison that $\Delta\lambda/\lambda \sim 1$ for thermal light sources). As was mentioned above, the intensity of a laser beam attains 10^8-10^{10} V/m.

The range of application of laser beams is extremely wide, from holography to the creation of pulsed thermonuclear reactor, from very fine instruments for ophthalmology to space communication.

Concluding this chapter, we note that sometimes lasers are considered to be objects of investigation of quantum macrophysics. Indeed, the laser effect can be regarded as a macroscopic quantum phenomenon associated with radiation in contrast to macroscopic quantum phenomena (like superfluidity and superconductivity) associated with a substance. It was mentioned more than once that photons are Bose particles, and quantum-mechanical rules allow an unlimited accumulation of identical Bose particles in the same state. A laser beam is a large (macroscopic) aggregate of identical photons.

Chapter 8
PLASMA

8.1. General Remarks

A partially or completely ionized gas forms a special state of aggregation known as **plasma.** In ordinary gases, the interaction between neutral atoms and molecules constituting them is manifested at the moments of collisions and rapidly attenuates with distance (the intermolecular potential varies with distance between molecules as r^{-6}). Since the plasma components (electrons and ions) are electrically charged, and the Coulomb potential decreases slowly (in proportion to r^{-1}), it is natural to expect that such a system will exhibit some collective properties. Indeed, all types of "electrically elastic" oscillations are readily excited in a plasma "constricted" as a whole by long-range Coulomb forces (plasma always produces a "noise").

Let us consider some properties of plasma using a completely ionized hydrogen plasma as an example. Such a plasma holds a promise of carrying out a controlled fusion reaction which offers a practically unlimited source of energy for future generations.

Fast fusion reactions of the explosive type have already been realized. In order to carry out a controlled fusion reaction with hydrogen isotopes converted into helium nuclei, a medium with a low density and a rather high temperature is required. A hydrogen plasma is suitable for this purpose. Indeed, in a high-temperature hydrogen plasma (i) hydrogen atoms are ionized (nuclei are bared); (ii) high-intensity nuclear collisions are ensured (reaction threshold); (iii) plasma can "easily" be isolated from the vacuum chamber walls with the help of magnetic fields.

8.2. Quantum Effects in Plasma. Tunneling of Nuclei Through a Potential Barrier

Let us estimate the "height" of electrostatic repulsive potential barrier emerging between two protons approaching each other to a short distance. Assuming that the minimum separation r between the two nuclei is of the order of 10^{-14} m, we obtain

$$U_{pp} = \frac{e^2}{r}, \quad U_{pp} \sim 10^{-14} \text{ J} \tag{8.2.1}$$

(it should be recalled that this formula is written in the CGS units). The corresponding temperature of the hydrogen plasma can be found from the relation

$$T \sim \frac{U_{pp}}{k}, \quad T \sim \frac{10^{-14}}{10^{-23}} \sim 10^9 \text{ K}. \tag{8.2.2}$$

At such a temperature, the energy of thermal motion of protons (actually, we must speak of hydrogen isotopes, i.e. deuterium ^{2}H and tritium ^{3}H) is sufficiently high to overcome the electrostatic repulsion and to form a compound system, i.e. a ^{4}He nucleus, by emitting a neutron n.

However, experiments show that the fusion reaction

$$^2\text{H} + {}^3\text{H} \rightarrow {}^4\text{He} + n \tag{8.2.3}$$

takes place even at a temperature $T \sim 10^7$ K. This means that hydrogen nuclei sometimes can somehow overcome the potential energy barrier at a kinetic energy lower by a factor of 100 than the barrier height.

In order to explain this paradox, we resort to quantum mechanics (see Chap. 4). Suppose that the real potential acting between a pair of protons approaching each other has a form similar to that shown in Fig. 8.1. In order to simplify the analysis, we go over to the square-topped potential barrier shown in Fig. 8.2*a*. We assume that a particle moves from left to right with an energy $E < U_0$. In region I, its probability wave has the form (Fig. 8.2*b*)

$$\psi_I = e^{ik_1 x} + Be^{-ik_1 x}, \tag{8.2.4}$$

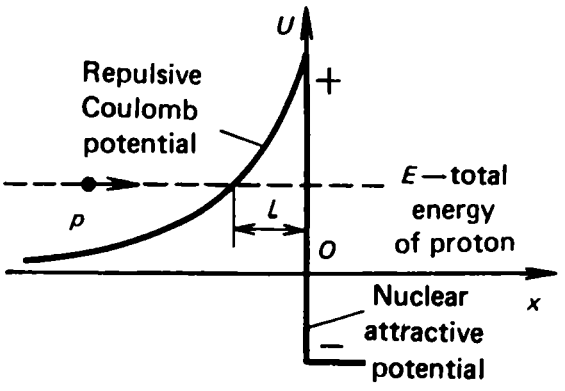


FIG. 8.1

where the first term describes the incident wave, and the second, the wave reflected from the barrier (B is a constant). Here,

$$k_1 = \frac{p}{\hbar} = \frac{1}{\hbar}\sqrt{2mE}. \tag{8.2.5}$$

In region II (when $E > U_0$), the probability wave

$$\psi_{II} = Ae^{ik_2x} \tag{8.2.6}$$

would describe an above-the-barrier transition, where

$$k_2 = \frac{1}{\hbar}\sqrt{2m(E-U_0)}. \tag{8.2.7}$$

However, in the case under consideration (when $E < U_0$), the quantity k_2 is imaginary, i.e. function (8.2.6) becomes attenuating:

$$\psi_{II} \sim e^{-k_2x}. \tag{8.2.8}$$

If the "tail" of wave (8.2.8) differs from zero at the second wall of the barrier, a passing wave with a small probability amplitude may exist in region III. This effect is known as the **under-the-barrier transition,** or **tunnel effect.**

In classical mechanics, a particle cannot be present in region II when $E < U_0$ for the obvious reason that in the case under consideration the kinetic energy would be negative and the velocity imaginary, which is meaningless.

In quantum mechanics, we can speak only of the total energy of a particle rather than of its kinetic and potential energy separately. As a matter of fact, the kinetic energy is determined by the value of velocity, while the potential

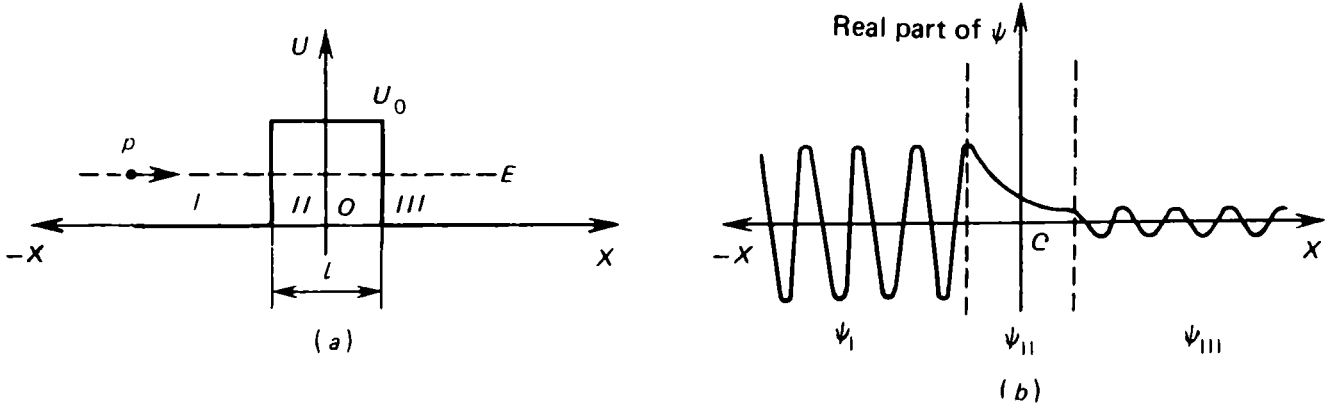


FIG. 8.2

energy by the coordinates. However, the velocity and the coordinate do not have definite values simultaneously. This circumstance makes the arguments formulated for a classical particle inapplicable to the quantum case.

There is no violation of the energy conservation law when a particle passes through a potential barrier since, according to the uncertainty relations (4.3.2), the localization of the particle in the barrier region makes the value of its energy completely indefinite.

Let us determine the probability w_{pp} of the under-the-barrier transition for two protons approaching each other in the presence of a strong Coulomb repulsion. This probability is determined by the squared modulus of (8.2.8), i.e.

$$w_{pp} \sim |\psi_{II}|^2 \sim e^{-2k_2 x}. \tag{8.2.9}$$

Let in this formula $x \sim l$ (see Fig. 8.2) and $k_2 = p/\hbar = 1/\lambda_p$, where λ_p is the length of the probability wave of a moving proton. Then (8.2.9) can be written in the form

$$w_{pp} \sim e^{-2l/\lambda_p}. \tag{8.2.10}$$

This formula shows that the probability of the under-the-barrier transition is noticeable when the barrier width l is of the order of the length λ of the probability wave of the particle.

Let us expand the exponent in (8.2.10):

$$2\frac{l}{\lambda_p} = 2\frac{l}{\hbar/p} = 2\frac{lmv}{\hbar} = 2\frac{lmv^2}{\hbar v}. \tag{8.2.11}$$

We assume that the energy $mv^2/2$ of a moving free proton is equal to half the height $U_{pp} = e^2/l$ of the potential barrier. Then (8.2.11) becomes

$$4\left(\frac{mv^2}{2}\right)\frac{l}{\hbar v} = 2\frac{e^2}{l}\frac{l}{\hbar v} = 2\frac{e^2}{\hbar v}.$$

As a result, (8.2.10) can be written in the form

$$w_{pp} \sim e^{-2e^2/(\hbar v)}. \tag{8.2.12}$$

This expression shows that in classical mechanics, when $\hbar \to 0$, the probability w_{pp} of the under-the-barrier transition vanishes. In quantum mechanics, for the case of hydrogen plasma under consideration, about one particle of sev-

eral hundred thousand particles will tunnel through the barrier.

Let us obtain this numerical estimate. Normally, the hydrogen plasma used in experiments has a density $(N_p/V) \sim 10^{19}$ protons/m^3, i.e. it is rarefied million times more than a gas under the normal atmospheric pressure. Protons are comparatively heavy particles ($m_p \sim 2 \times 10^3 m_e \simeq 1.7 \times 10^{-27}$ kg). According to these data, the degeneracy temperature (see Eq. (5.8.10)) of the hydrogen plasma is

$$T_d \sim \frac{\hbar^2}{m_p k} \left(\frac{N_p}{V}\right)^{2/3},$$

$$T_d \sim \frac{(10^{-34})^2 \times (10^{19})^{2/3}}{1.7 \times 10^{-27} \times 1.4 \times 10^{-23}} \text{ K} \sim 10^{-5} \text{ K}. \tag{8.2.13}$$

In other words, the high-temperature hydrogen plasma obeys the laws of classical statistics. Therefore, the mean thermal velocity $\langle v_p \rangle$ of protons in the plasma can be determined from formula (5.8.7):

$$\langle v_p \rangle = \sqrt{\frac{3kT}{m_p}} \sim \sqrt{\frac{3 \times 1.4 \times 10^{-23} \times 10^7}{1.7 \times 10^{-27}}} \text{ m/s}$$

$$= 5 \times 10^5 \text{ m/s}. \tag{8.2.14}$$

Considering that the exponent in formula (8.2.12) is expressed in the CGS units, we obtain

$$2 \frac{e^2}{\hbar v} \sim 10,$$

i.e. the probability is given by

$$w_{pp} \sim e^{-10} = 4.5 \times 10^{-5}.$$

Let us estimate the total number of protons participating in the fusion reaction at $T \sim 10^7$ K (calculations are carried out for the plasma volume of 1 m^3). Before that, we calculate the fraction of the total number of protons having the mean thermal velocities $\langle v_p \rangle \sim 5 \times 10^5$ m/s at $T \sim 10^7$ K. This fraction is determined from the relation

$$\left(\frac{N_p}{V}\right) w(v), \tag{8.2.15}$$

where

$$w(v) \sim e^{-mv^2/(2kT)} \tag{8.2.16}$$

is just the equilibrium statistical velocity distribution of protons (see formula (5.8.5) and Fig. 5.12). The numerical solution of Eq. (8.2.16) can be obtained by carrying out consecutively the following operations:

$$\frac{mv^2}{2kT} \sim \frac{1.7 \times 10^{-27} (5 \times 10^5)^2}{2 \times 1.4 \times 10^{-23} \times 10^7} \simeq 1.5,$$

and $w(v) \sim e^{-1.5} = 0.2$, i.e. the fraction of protons having the required thermal velocities is about 20%. Finally, the total number of protons that undergo elementary acts of nuclear fusion can be written in the form

$$\left(\frac{N_p}{V}\right) w(v)\, w_{pp}, \tag{8.2.17}$$

and its numerical value is

$$10^{19} \times 0.2 \times 4.5 \times 10^{-5} \sim 10^{14}.$$

8.3. Relativistic Effects in Plasma. Mass Defect in Nuclear Fusion and Energy Liberated in the Process

In order to calculate the value of the energy liberated in a fusion reaction, we must consider the most general relativistic relations. Indeed, a nucleus is not only a quantum-mechanical but also a relativistic system. This can easily be verified by using the quantum-mechanical uncertainty principle

$$\Delta p \sim \frac{\hbar}{\Delta r} \sim \frac{10^{-34}}{10^{-15}} \text{ kg·m·s}^{-1} \sim 10^{-19} \text{ kg·m·s}^{-1}$$

and the velocity-momentum relation

$$v_N \sim \frac{p}{m_N} \sim \frac{10^{-19}}{10^{-27}} \text{ m/s} \sim 10^8 \text{ m/s}.$$

Here, m_N is the mass of a nucleon, and v_N is the "velocity" of the nucleon in the nucleus, which is of the order of the speed of light c.

Let us apply the results obtained in Sec. 2.7 to fusion reactions. We shall consider a simple example of formation of

a deuterium (^{2}H) nucleus from a free proton p and neutron n:

$$p + n \rightarrow {}^2\mathrm{H}.$$

We shall write the measured values of the masses of these particles:

$$m_p = 1.67243 \times 10^{-27} \text{ kg},$$

$$m_n = 1.67474 \times 10^{-27} \text{ kg},$$

$$M_{^2\mathrm{H}} = 3.34321 \times 10^{-27} \text{ kg}.$$

Let us calculate the mass defect by formula (2.7.11):

$$\Delta M = M_{^2\mathrm{H}} - (m_p + m_n)$$

$$= (3.34321 - 3.34717) \times 10^{-27} \text{ kg} = -0.00396 \times 10^{-27} \text{ kg}$$

$$= -3.96 \times 10^{-30} \text{ kg}.$$

We can now calculate the corresponding binding energy for a deuterium nucleus (see Eq. (2.7.12)):

$$\Delta E = c^2 \Delta M = (3 \times 10^8)^2 \times 3.96 \times 10^{-30} \text{ J}$$

$$= 3.564 \times 10^{-13} \text{ J} = 2.225 \text{ MeV}.$$

This energy is liberated in the fusion of a deuterium nucleus.

It should be noted that the energy yield of a reaction of the type (8.2.3) is 17.6 MeV. Almost the entire energy is converted into the kinetic energy of the liberated neutron.

8.4. Plasma Statistics. Equation of State for Plasma. Thermal Radiation of Plasma

Let us consider some elementary processes in a plasma which will be treated as a large aggregate of particles. What must be the statistics describing a hot rarefied plasma?

General methods of statistical physics (see Sec. 5.8.2) allow us to answer this question unambiguously. Indeed, for this purpose, we must only estimate the degeneracy temperature T_d of the system (see the end of Sec. 5.8.3). If the temperature of the system $T \gg T_d$, it obeys the laws of classical statistics. If, however, $T \ll T_d$, the system obeys the laws of quantum statistics. For a hydrogen plasma consisting of two (nuclear and electron) subsystems, the degener-

acy temperature should be estimated for the two subsystems separately. For the nuclear component, the estimate of T_d has already been made (see Eq. (8.2.13)). For the electron component, we have

$$T_d \sim \frac{\hbar^2}{m_e k}\left(\frac{N_e}{V}\right)^{2/3},$$

$$T_d \sim \frac{(10^{-34})^2 \times (10^{19})^{2/3}}{10^{-30} \times 10^{-23}}\ \mathrm{K} \sim 10^{-2}\ \mathrm{K}.$$

Therefore, we find that a hot hydrogen plasma has $T \gg T_d$. Thus, the collective properties of plasma can be described well enough by the methods of classical statistics.

While estimating in Sec. 8.2 the number of plasma particles having preset velocities, we used the classical equilibrium Maxwell distribution (8.2.16). While calculating the mean thermal velocity of plasma particles, we used formula (8.2.14) of classical statistics.

Let us derive an expression for the thermal energy of plasma and its thermodynamic equation of state. Is it possible to use the well-known relations for the classical ideal gas for this purpose? The answer is affirmative if the plasma density is so low that the energy of Coulomb interaction between the plasma particles is small in comparison with their kinetic energy, i.e.

$$\frac{e^2}{r} \ll kT. \tag{8.4.1}$$

Taking into account the fact that the mean distance between the plasma particles is

$$r \sim n^{-1/3}, \tag{8.4.2}$$

where n is the number density of particles, and substituting (8.4.2) into (8.4.1), we obtain

$$\frac{e^2}{n^{-1/3}} \ll kT,$$

or

$$n \ll \left(\frac{kT}{e^2}\right)^3. \tag{8.4.3}$$

This criterion must be satisfied by the density of plasma that can be regarded as the ideal gas of particles.

For the hydrogen plasma under consideration, $n \sim 10^{19}$ particles/m^3, and the right-hand side of (8.4.3) is of the order of

$$\left(\frac{kT}{e^2}\right)^3 \simeq 10^{30}$$

(it should be recalled that the expression is written in the CGS units). Thus, inequality (8.4.3) is satisfied with a large margin. Therefore, a hot rarefied plasma can be treated as an ideal gas, i.e. the total thermal energy E of the plasma is reduced to the sum of the kinetic energies of individual particles:

$$E = (N_p + N_e)\, kT. \tag{8.4.4}$$

The thermal pressure of the plasma is given by

$$p \approx \frac{E}{V} = (n_p + n_e)\, kT. \tag{8.4.5}$$

Plasma radiation plays a significant role in a hot plasma. Let us try to estimate the order of magnitude of the temperature at which the radiant energy density $(\sigma/c)\, T^4$ (see Eq. (5.9.16)) and the thermal energy density of a substance (see Eq. (8.4.5)) become equal. We write the expression

$$\frac{\sigma}{c}\, T^4 \approx nkT, \tag{8.4.6}$$

where σ is the Stefan-Boltzmann constant. Substituting the numerical values into (8.4.6), we obtain

$$\frac{5.7 \times 10^{-8}}{3 \times 10^8}\, T^4 \simeq 10^{19} \times 10^{-23}\, T,$$

or

$$10^{-16}\, T^4 \simeq 10^{-4}\, T. \tag{8.4.7}$$

This equation is satisfied for $T \simeq 10^4$ K (the reader can easily verify this).

For such a temperature, the hydrogen plasma is ionized only partially. Both the ionization of atoms

$$\mathrm{H} \rightarrow p + e^-$$

(since at $T \simeq 10^4$ K hydrogen molecules have already been dissociated) and recombination processes

$$p + e^- \rightarrow \mathrm{H}$$

occur in such a plasma in which atoms can be in the ground as well as in various excited states. We could determine the equilibrium concentrations of all the components of such a cold plasma. Here, the situation is as follows: (i) excited atoms emit a discrete radiation (see Sec. 4.7.2); (ii) when a hydrogen nucleus meets an electron and is combined with it to form an atom, the excess energy is emitted; (iii) free nuclei and electrons coil around the lines of force of an external stabilizing magnetic field (see below) and emit a classical continuous radiation (see Secs. 3.6.2 and 3.7.3).

As we go over to a hot plasma consisting of only nuclear and electron ideal gases, from all types of radiation (discrete, recombination, and continuous) we will be left only with the classical electromagnetic radiation having a continuous spectrum and associated with the accelerated motion of free charged particles.

At $T \sim 10^7$ K, the intrinsic radiation of a plasma lies in the X-ray region. This can easily be proved if we use Eq. (5.9.7):

$$\lambda = \frac{2\pi\hbar c}{4.96kT},$$

$$\lambda \simeq \frac{6.3 \times 10^{-34} \times 3 \times 10^8}{5 \times 1.4 \times 10^{-23} \times 10^7}\ \mathrm{m} \sim 10^{-10}\ \mathrm{m}.$$

For a number density of 10^{19} particles/m^3, the plasma is transparent for its radiation which is not in thermal equilibrium with the substance. For this reason, relation (8.4.7) can be applied to an analysis of the situation only conditionally. It is only clear that two competing tendencies should be observed in the system. One of them is due to the energy of the substance which grows in proportion to temperature, while the other is associated with the radiant energy which grows in proportion to the fourth power of temperature. When the temperature is high enough, practically the entire energy supplied to the system is carried away by radiation, and the substance remains "cold". However, we must also take into account the fact that for $T \sim 10^7$ K, the latent

energy of the substance comes into play (relativistic rest energy of particles), which is (partially) released in fusion reactions. Its contribution violates the energy balance of the system to such an extent that the fusion reaction becomes self-sustained.

8.5. Plasma Kinetics. Mobility of Ions and Its Relation with Diffusion. Electrical Conductivity of Plasma

In gaseous media, kinetic transport phenomena boil down to the processes of viscosity, heat conduction, and diffusion (see Chap. 6). In plasma, electrical conduction is present along with these phenomena.

Let us consider a not-too-rarefied plasma for which the kinetic coefficients still have their direct meaning (see Sec. 6.5). First of all, we are interested in the electrical conduction of the plasma. It turns out that the electrical conduction of the plasma is connected with diffusion of plasma particles.

In an external electric field **E**, the random thermal motion of ions is superimposed on their slow displacement in the direction of the field, known as **drift**. The drift velocity $\mathbf{v}_{dr}$ of charged particles along the electric field can be written, by analogy with Eq. (7.2.4), in the form

$$\mathbf{v}_{dr} \sim \frac{\tau}{m} \mathbf{F}, \tag{8.5.1}$$

where

$$\mathbf{F} = e\mathbf{E}. \tag{8.5.2}$$

The proportionality factor between the drift velocity and the constant force in (8.5.1) is usually denoted by

$$b = \frac{\tau}{m} \tag{8.5.3}$$

and called **mobility**. For light particles in a rarefied medium, the mobility is high. Let us now use relation (6.5.2) for the diffusion coefficient D. We express D in terms of the mean time τ between collisions of particles according to Eq. (6.5.1*) and take into account Eq. (8.5.3). This gives

$$D \sim vl \sim v^2\tau \sim mv^2 b. \tag{8.5.4}$$

If we note that in this formula, according to Eq. (5.8.7), $mv^2 \sim kT$, we finally obtain

$$D = kTb. \tag{8.5.5}$$

It is remarkable that the obtained relation between the mobility of particles and their diffusion in the medium is exact.

Let us now consider the electrical conductivity of a plasma. The initial quantitative relation here is (7.2.3):

$$\mathbf{j}_e = en_e\mathbf{v}_{dr}, \tag{8.5.6}$$

where $\mathbf{j}_e$ is the current density of the electron component of the plasma. According to Eq. (8.5.3), the mobility b_1 of the ion component is lower by three orders of magnitude. Using Eqs. (8.5.1)-(8.5.3), we can write (8.5.6) in the form

$$\mathbf{j}_e = e^2n_eb_e\mathbf{E}. \tag{8.5.7}$$

Assuming that Ohm's law (7.2.2) is satisfied for the electrical conductivity of the plasma, we obtain

$$\sigma = e^2n_eb_e. \tag{8.5.8}$$

Using Eq. (8.5.5), we get

$$\sigma = \frac{e^2n_eD}{kT}. \tag{8.5.9}$$

For $T \sim 10^7$ K, the electrical conductivity σ of the hydrogen plasma is very high and attains the values of that for solid metals (like copper).

8.6. Magnetohydrodynamics and Plasma Instabilities. Tokamaks

If the number density n of plasma particles is such that the mean free path l of the particles is small in comparison with the characteristic size L of the system, in the analysis of transport phenomena in plasma we can use an approach in which plasma is described as a continuum (see the end of Sec. 6.2). It would appear that the methods of ordinary hydrodynamics can be successfully used for describing nonequilibrium states of a dense plasma. However, the dense plasma is a "conducting liquid", and its motion, say, in an

external magnetic field considerably differs from the normal liquid flow. Indeed, if (i) the plasma flows in a constant magnetic field, its charged particles experience the action of the electrodynamic Lorentz force, and (ii) an external varying magnetic field induces currents in the plasma, which themselves produce an intrinsic magnetic field which affects the motion of the plasma. Therefore, the dense plasma should be described by a combined system of equations of hydrodynamics and electrodynamics. This forms the essence of the **magnetohydrodynamics.**

What are the main results of the application of magnetohydrodynamics to the description of plasma? Since, as was mentioned above, the electrical conductivity of a plasma can be high enough, it is natural to introduce the model of a perfectly conducting ($\sigma = \infty$) liquid. In this case, an external magnetic field cannot penetrate the plasma since otherwise infinitely large currents would be induced in it. Therefore, the external magnetic field must exert a certain pressure on the plasma. The expression for magnetic pressure can easily be written if we use formula (3.8.8) for the energy density of a magnetic field:

$$p \approx \frac{E}{V} = \frac{H^2}{8\pi}. \tag{8.6.1}$$

Let us consider the effect of a **linear pinch discharge.** If an electric discharge takes place in a gas-filled chamber, firstly, the gas is ionized and plasma is formed as a result of Joule heat losses, and, secondly, the intrinsic magnetic field of the discharge current separates the formed plasma from the chamber walls and constricts it into a thin column (pinch) (Fig. 8.3 illustrates this effect). Pinching of a plasma is possible if the magnetic pressure is comparable in the order

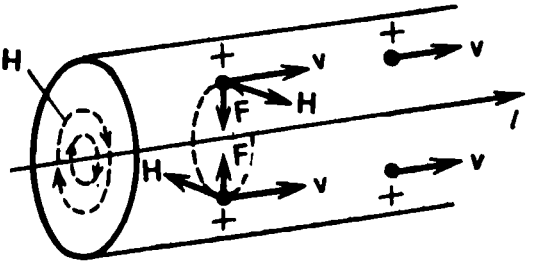


FIG. 8.3

of magnitude with the thermal pressure of the plasma, i.e. when

$$2nkT = \frac{H_I^2}{8\pi}. \tag{8.6.2}$$

Here, we have used expressions (8.6.1) and (8.4.5). The magnetic field strength H_I of a current is given by(see Eq. (3.5.2))

$$H_I = \frac{2I}{cr_0}, \tag{8.6.3}$$

where I is the current, and r_0 is the radius of the plasma column. Substituting (8.6.3) into (8.6.2), we obtain

$$I^2 = 4c^2\pi r_0^2 nkT. \tag{8.6.4}$$

In most of plasma experiments, $I \sim 10^5$ A (1 A $= 3 \times 10^9$ CGS units), $r_0 \simeq 0.1$ m, and $n \sim 10^{19}$ m^{-3}. Then the temperature of the plasma column is

$$T = \frac{I^2}{4c^2\pi r_0^2 nk}.$$

Calculation in this formula is made in the CGS units. As a result, we obtain

$$T = 5 \times 10^7 \text{ K}.$$

To the astonishment of researchers, the plasma column existed for a very short time ($\sim 10^{-6}$ s). The reason behind its failure was that the plasma column was in a state of unstable equilibrium. A small external perturbation (like bending or pinch of the plasma column, see Fig. 8.4) led to

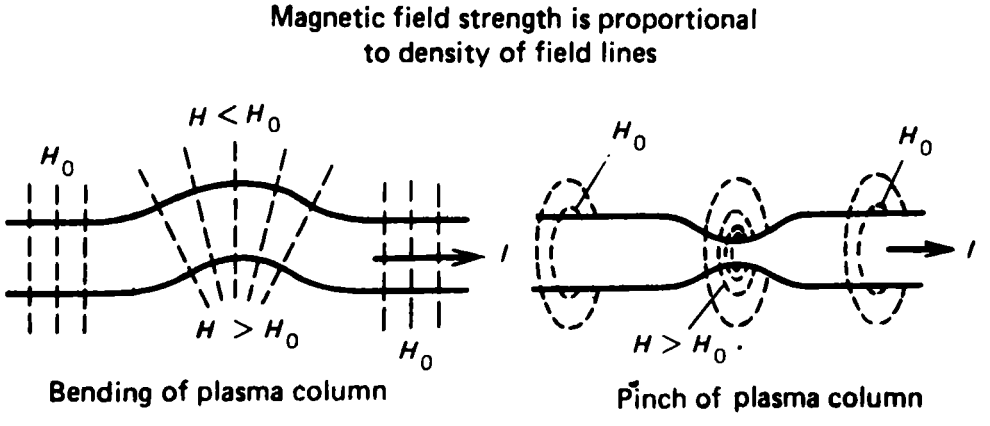


FIG. 8.4

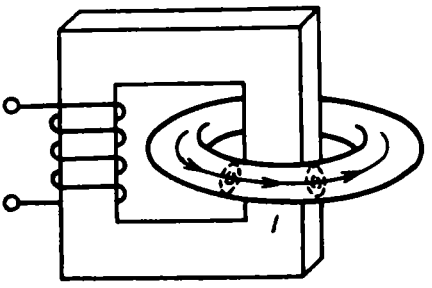


FIG. 8.5

a local change in the intrinsic magnetic field of the plasma current (and hence in the magnetic pressure) which enhanced the deviation from the equilibrium configuration. A strong longitudinal magnetic field can be effectively used to stabilize the plasma column. The plasma confinement time considerably increases in this case.

At the modern stage of plasma research, the most popular are the experimental plasma devices of the Tokamak type. In fact, a Tokamak is a transformer in which the secondary winding consists of a single turn (Fig. 8.5). This turn is made in the form of a closed annular chamber like a torus (doughnut). The annular chamber is filled with a deuterium-tritium gas mixture under a low pressure. As the current is passed through the primary winding, a gas breakdown takes place in the chamber. The gas is ionized, and the current flowing through it heats the gas to a high temperature. A magnetic field used for the plasma confinement and stabilization is produced with the help of coils arranged along the torus.

A developed fusion reaction has already been obtained on the Soviet device Tokamak-10 with a plasma confinement time of 0.1 s.

8.7. Oscillations and Waves in Plasma. Propagation of Radio Waves in the Ionosphere

Since plasma is a complex electrodynamic and hydrodynamic system, oscillations can be excited in it, and waves of different types may accordingly propagate. Oscillations in a plasma can be of electromagnetic, electrostatic, magnetohydrodynamic, and acoustic origin. In the general case, all these types of oscillations are always present in a system.

In the absence of an external magnetic field, acoustic and electrostatic oscillations can take place in a plasma.

Indeed, since a plasma contains two groups of particles with opposite charges, the plasma is electrically neutral as a whole, and

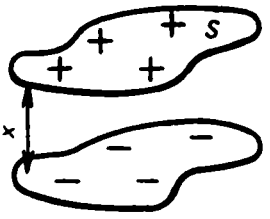


FIG. 8.6

(i) a joint displacement of particles of both signs in the same direction leads to the emergence of a local region of elevated pressure, which generates elastic (acoustic) vibrations;

(ii) a displacement of unlike particles in opposite directions (their separation) leads to the emergence of a local electric field which causes oscillations of charges. Clearly, the corresponding plasma waves can only be longitudinal unlike electromagnetic waves which are transverse (see Eqs. (3.7.15) and (3.8.1)).

Let us derive the quantitative expression for plasma oscillations. We assume that a local separation of charges in a plasma occurs in a certain plane layer (Fig. 8.6). Then it is natural to apply the formula for a parallel-plate capacitor (7.7.48):

$$C = \frac{S}{4\pi x}, \tag{8.7.1}$$

where S is the area of the surface, x is the separation between the planes, and C is the capacitance of the "capacitor". The charge on the plates of the plasma capacitor is

$$Q = enSx. \tag{8.7.2}$$

The definition of capacitance (7.5.8) implies that the potential difference across the layer is

$$\Delta\varphi = \frac{Q}{C} = 4\pi enx^2. \tag{8.7.3}$$

Here, we have used expressions (8.7.2) and (8.7.1). Knowing $\Delta\varphi$ and the linear size x of the region of charge separation, according to Eq. (3.4.14), we can find the electric field

$$E = -\frac{\Delta\varphi}{x} = -4\pi enx. \tag{8.7.4}$$

In accordance with Eq. (3.6.1), this field imparts to each charge the acceleration

$$\frac{d^2x}{dt^2} = \frac{eE}{m} = -\frac{4\pi e^2 n}{m}\, x. \tag{8.7.5}$$

We have obtained an equation of the type (7.7.26) or, to be more precise, (7.7.28). Similar equations must have similar solutions. Consequently,

$x = \cos \omega t$,

where

$$\omega \equiv \omega_p = \sqrt{\frac{4\pi e^2 n}{m}}. \tag{8.7.6}$$

In view of the high mobility of electrons, we deal with electron plasma oscillations. Then the frequency ω_p is determined only by the number density n_e of the plasma.

Let us consider the passage of electromagnetic waves (in particular, radio waves) through a plasma. While writing expression (7.7.64) for the refractive index in metals, we made the only assumption about the presence of free electrons in the medium. But this assumption is also satisfied for plasma, and hence all the relations contained in Sec. 7.7.7 are valid for it. In particular, formula (7.7.65)

$$n^2 = 1 - \frac{\omega_p^2}{\omega^2} \tag{8.7.7}$$

with the plasma frequency ω_p (cf. Eq. (7.7.66) with (8.7.6)) is fundamental in the analysis of propagation of radio waves through a plasma. For example, for $\omega^2 < \omega_p^2$, the refractive index (8.7.7) becomes imaginary, i.e. there is no wave in the plasma (it is either absorbed or reflected). For $\omega \gg \omega_p$, the refractive index is real-valued and differs from unity only slightly. This means that the plasma is transparent to radiation at such a frequency.

Let us consider the ionospheric layer of the Earth as an object of application of the results obtained above. The ionosphere is a plasma formation with a complex composition. The reason behind its ionization is the ultraviolet solar radiation, as well as flows of fast charged particles emitted by the Sun. The maximum electron number density of the ionosphere determined with the help of probes is 10^{12} electrons/m^3. These data make it possible to calculate the eigenfrequencies of plasma oscillations in the ionosphere:

$$\omega_p = \sqrt{\frac{4\pi e^2 n}{m}}.$$

Substituting the numerical values expressed in the CGS units, we obtain

$\omega_p \simeq 6 \times 10^7 \ s^{-1}$.

Let us calculate the wavelength of electromagnetic waves whose frequency coincides with the plasma frequency in the ionosphere. By using formula (3.8.6), we obtain

$$\lambda = \frac{2\pi c}{\omega} = \frac{6.28 \times 3 \times 10^8}{6 \times 10^7} \text{ m} \simeq 30 \text{ m}.$$

Therefore, radio waves with $\lambda > 30$ m will be reflected from the ionosphere as from a metal mirror. The communication with satellites and orbital stations requires waves much shorter than 30 m.

We hope that the reader has got a correct idea that the general regularities of propagation of electromagnetic waves in metals, plasma, the ionosphere of the Earth, and the atmosphere of stars are the same, the only difference being in the values of n_e and hence of ω_p.

Chapter 9
STELLAR AND PRESTELLAR STATES OF MATTER

9.1. State of Matter at Ultrahigh Temperatures and Densities

Along with the investigations of the normal states of matter, the study of substances at ultrahigh temperatures and densities is of great interest.

Let us analyze qualitatively the change in the states of a substance with a gradual increase in temperature (assuming that the density remains not very high). It was shown that in this process the condensed state of a substance (solid and liquid) is transformed to the gaseous state. Ordinary gases are molecular gases. An increase in temperature to several thousand kelvins leads to **thermal dissociation** (decomposition of molecules into atoms constituting them), and the gases become atomic.

At temperatures of the order of 10^4 K, atoms constituting a gas are ionized. The ionized substance (plasma) consists only of ions and electrons ($T \sim 10^6$ K).

At temperatures of the order of 10^7 K, the plasma is completely ionized, and the substance consists of "bare" nuclei and free electrons. A further increase in temperature leads to nuclear transformations ($T \sim 10^8$ K).

At temperatures above 10^9 K, nuclei are decomposed, and the substance consists only of protons and electrons[1] ($T \sim 10^{11}$ K).

Finally, when the temperature rises above 10^{13} K, a large-scale transformation of elementary particles becomes possible. For example, the energy required for the formation of nucleon-antinucleon pairs is of the order of $m_N c^2$, where m_N is the mass of a nucleon. The above value of temperature is obtained from the relation $m_N c^2 \sim kT$.

---

[1] A neutron is an unstable particle which decays into a proton, an electron, and an antineutrino. The latter practically does not interact with the substance and leaves the system.

When we considered above the change in the state of a substance along the "temperature axis", the pressure was assumed to remain normal. Let us now analyze qualitatively the change in the state of a substance with increasing density, assuming that the temperature remains at the normal level.

In a substance compressed to about 10^8 atm, the electron shells of atoms are deformed, and their internal energy sharply increases. The electric fields of individual atoms are superimposed, and as a result, the bond of electrons of atomic shells to a certain atom becomes weaker and weaker. The free motion of outer electrons becomes possible ("metallization" of the substance).

Under moderate pressures, substances exhibit diverse properties and reveal a strong and nonmonotonic dependence on the chemical composition, while in compressed matter, there is a tendency to "smooth" the individual properties. This effect can easily be explained. Indeed, the outer electron shells of atoms in a substance which determine the above-mentioned nonmonotonicity cease to exist at pressures of the order of 10^8 atm since the electrons constituting them are detached from the atoms. The inner electron shells become denser, and the electron number density in them is comparatively uniform.

Thus, at a high compression of a substance ($\sim 10^{12}$ atm), the role of interaction of the electrons of its atoms with nuclei becomes insignificant, and the substance can be regarded as a high-density electron gas (such a gas is referred to as degenerate).

When the density and pressure of the gas become of the order of 10^9 kg/m^3 and 10^{18} atm respectively,[2] the electron gas becomes relativistic (i.e. the mean energy of an electron becomes comparable with $m_e c^2$).

A further increase in density leads to states in which nuclear reactions turn out to be thermodynamically advantageous. These involve capturing of electrons by nuclei and simultaneous emission of a neutrino. As a result of such a reaction, the charge of a nucleus is reduced (at a constant

---

[2] For the sake of comparison, it should be mentioned that one of the most dense substances like osmium has a density of 2.2×10^4 kg/m^3 at normal temperature and a pressure of 1 atm.

mass), which generally leads to a decrease in the binding energy of the nucleus.

At still higher densities and pressures, a further capture of electrons by nuclei takes place, accompanied by a further decrease in the nuclear charge. Ultimately, nuclei containing too many neutrons become unstable and disintegrate. At densities of about 10^{14} kg/m^3 and pressures of about 10^{24} atm, the number of neutrons exceeds the number of electrons, and at densities of the order of 10^{15} kg/m^3, the neutron pressure exceeds the electron pressure. In this density range, the substance can be treated mainly as a neutron Fermi gas (naturally, a certain amount of protons and electrons formed as a result of neutron decay is always present in this gas).

At pressures of the order of 10^{27} atm, the neutron gas has the same density as the nuclear matter, i.e. about 10^{17} kg/m^3. This state of matter can exist up to temperatures of the order of 10^{12} K.

9.2. Stars as Gaseous Spheres

9.2.1. Calculation of Pressure and Temperature at the Centre of a Star. The matter in stars exists at ultrahigh temperatures and densities. The main characteristics of stars include their mass M_* and radius R_*. We shall not consider the methods of determining these quantities empirically.

How can we estimate the pressure p at the centre of a star? We shall use the method of dimensional analysis. It is natural to express the required pressure in terms of the main characteristics of the star, viz. its mass M_* and radius R_*. Of course, the gravitational constant G must also appear in the expression. The combination

$$p \sim G^x \cdot M_*^y \cdot R_*^z \qquad (*)$$

of the quantities under consideration is unique. It contains the unknowns x, y, and z which have to be determined. For this purpose, we write from relation (*) the dimensions of the participating quantities:

$$[p] = [G]^x \cdot [M_*]^y \cdot [R_*]^z,$$

or

$$\mathrm{kg \cdot m^{-1} \cdot s^{-2}} = (\mathrm{kg^{-1} \cdot m^3 \cdot s^{-2}})^x \cdot (\mathrm{kg})^y \cdot (\mathrm{m})^z = \mathrm{kg^{-x+y} \cdot m^{3x+z} \cdot s^{-2x}},$$

Comparing the exponents of both sides of this equation, we obtain the following system of equations:

$$\left.\begin{aligned} -x + y &= 1, \\ 3x + z &= -1, \\ -2x &= -2. \end{aligned}\right\}$$

Solving this system, we obtain $x = 1$, $y = 2$, and $z = -4$. Therefore,

$$p \sim G \cdot M_*^2 \cdot R_*^{-4},$$

$$p \sim \frac{GM_*^2}{R_*^4}. \tag{9.2.1}$$

In this expression, the pressure at the centre of the star is expressed in terms of the quantities determined in laboratory experiments (G) and by astronomical observations (M_* and R_*).

Let us determine the order of magnitude of pressure, for example, at the centre of the Sun, for which $M_\odot = 2 \times 10^{30}$ kg and $R_\odot = 7 \times 10^8$ m:

$$p_\odot \sim \frac{6.7 \times 10^{-11} (2 \times 10^{30})^2}{(7 \times 10^8)^4} \text{ Pa} \simeq 1.1 \times 10^{15} \text{ Pa} \simeq 10^{10} \text{ atm}.$$

Although most stars are undoubtedly plasma formations, we shall model a star in the form of a gaseous sphere for which the equatin of state is

$$pV = \frac{m}{M} RT, \tag{9.2.2}$$

where R is the gas constant, and M is the molar mass.

For a stationary star, the gravitational compression must be compensated by the thermal counterpressure:

$$p_{\text{grav}} \approx p_{\text{therm}}. \tag{9.2.3}$$

Substituting (9.2.1) and (9.2.2) into this expression, we obtain

$$G \frac{M_*^2}{R_*^4} \approx \frac{\rho}{M} RT_c. \tag{9.2.4}$$

The mean density of the star can be estimated as

$$\rho \sim \frac{M_*}{R_*^3}. \tag{9.2.5}$$

Using (9.2.5) in (9.2.4), we can write the expression for the temperature at the centre of the star:

$$T_c \sim G\frac{M_* M}{R_* R}. \tag{9.2.6}$$

For the Sun, $M = 1 \times 10^{-3}$ (hydrogen star), and the temperature at the centre of the Sun is

$$T_{\odot c} \sim \frac{6.7\times10^{-11}\times2\times10^{30}\times1\times10^{-3}}{7\times10^{8}\times8.3}\ \text{K} \sim 2\times10^{7}\ \text{K}.$$

9.2.2. Temperature of the Surface and the Total Emissive Power of a Star. The radiation from stars can be treated as equilibrium thermal radiation. Then the results obtained in Sec. 5.9 are completely applicable to it. In this section, we calculated the temperature of the outer layers of the Sun with the help of formula

$$T = \frac{0.29}{\lambda_{max}}.$$

This formula is written in the CGS units. Calculations give $T_\odot \simeq 5300$ K. Here, we take into account the fact that the maximum of the energy emitted by the Sun corresponds to the yellow-green region of the spectrum.

Knowing the temperature of emitting layers, we can determine the radiation intensity

$$j = \sigma T^4,$$

i.e. the power emitted by a unit surface. The total emissive power of the star, referred to as its **luminosity**, is given by

$$J = 4\pi R_*^2 \sigma T^4. \tag{9.2.7}$$

For the Sun, the luminosity is

$$J_\odot = 4\pi R_\odot^2 \sigma T_\odot^4$$
$$= 4\times3.14\,(7\times10^8)^2\times5.7\times10^{-8}\,(5.3\times10^3)^4\ \text{W} = 2.8\times10^{26}\ \text{W}.$$

9.2.3. Energy Transfer in Stars. Calculations made for the Sun show that its temperature of the surface layers is $T_{sur} \sim$

(5-6) $\times 10^3$ K and $T_c \sim 2 \times 10^7$ K at the centre. Such a temperature difference indicates that intense heat transfer takes place in the Sun.

There are three mechanisms of heat transfer: conduction, convection, and radiation. For ordinary stars, the role of heat conduction is very small, while for white dwarfs, it is quite significant (see below). As regards convection, for example, in the Sun, it is significant only for surface layers. Thus, radiation plays the decisive role in the energy transfer in stars.

A stellar matter is opaque to its own radiation. The radiation is absorbed and re-emitted by the upper layers. The mechanisms of absorption are: (i) photoionization of light atoms; (ii) absorption of light by electron shells of heavy elements; (iii) scattering of light by free electrons.

The absorption of light can be quantitatively analyzed as follows. A column of substance is mentally cut from the star along its radius so that it extends from the centre of the star to its surface. The cross-sectional area of this column is 1 m^2, the length coincides with the star radius R_*, while the density of the column of substance is equal to the mean density ρ of the star. Then for the column of substance under consideration, we have

$$\text{Fraction of absorbed light} = \gamma\rho R_*,$$

where γ is the opacity (for the Sun, $\gamma = 2$ m^2/kg). The intensity of radiation passing through the column is

$$j \equiv \frac{J}{4\pi R_*^2} = \frac{\sigma T_c^4}{\gamma\rho R_*}.$$

Substituting relations (9.2.6) for T_c and (9.2.5) for ρ into this expression, we obtain

$$J \sim \frac{4\pi\sigma G^4 M^4}{\gamma R^4} M_*^3. \qquad (9.2.8)$$

This is the well-known **mass-luminosity relation** which is observed for a huge number of stars.

9.3. Sources of Stellar Energy

9.3.1. Analysis of Possible Sources of Stellar Energy. In the model of a star as a hot gaseous sphere constricted by gravitational forces, the intrinsic gravitational energy may

be a source of energy. As the energy is emitted and the star cools, the star becomes compressed, and this leads to a new increase in temperature. Formally, this process can be continued until the star contracts to a point.

How long can this mechanism ensure the energy liberation by a star? In order to answer this question, we must compare the intrinsic potential gravitational energy of the star

$$U = G\frac{M_*^2}{R_*}$$

with its emissive power J. For the Sun, we have

$$U_\odot \sim \frac{6.7\times10^{-11}\,(2\times10^{30})^2}{7\times10^8}\ \text{J} \sim 4\times10^{41}\ \text{J}$$

and

$$J_\odot \sim 2.8\times10^{26}\ \text{W}.$$

This gives

$$\frac{U_\odot}{J_\odot} \sim \frac{4\times10^{41}\ \text{J}}{3\times10^{26}\ \text{W}} \sim 10^{15}\ \text{s} \sim 3\times10^7\ \text{years}.^{3}$$

This time period is too short since the Sun exists for about 10^9 years. Thus the observed duration of the active life of a star cannot be explained by gravitational compression.

Let us now consider fusion reactions as a source of stellar energy. The reactions of proton-proton type are the most likely candidates for this role. They ultimately result in the transformation of four nucleons into a helium nucleus:

$$4N \rightarrow {}^4\text{He}.$$

Let us now make calculations. We shall use the atomic mass units (amu)

$$1\ \text{amu} = \frac{1}{12}\,m_{^{12}\text{C}} = 1.66\times10^{-27}\ \text{kg}.$$

In these units,

$$m_N = 1.00813\ \text{amu}, \quad 4m_N = 4.03252\ \text{amu},$$
$$m_{^4\text{He}} = 4.00389\ \text{amu}.$$

---

[3] On the average, 1 year $= 3\times10^7$ s.

The mass defect of the reaction is

$$\Delta m = m_{^4\mathrm{He}} - 4m_N = -0.02863 \text{ amu},$$

and the corresponding liberated energy is

$$\Delta E = c^2 \Delta m = (3 \times 10^8)^2 \times 0.02863 \times 1.66 \times 10^{-27} \text{ J}$$
$$= 4.3 \times 10^{-12} \text{ J}.$$

Thus, in the reaction under consideration, the fraction $\Delta m/(4m_N) = 0.007$ of the total mass is transformed into energy.

Let us now consider the star as a whole. For the Sun (under the assumption that it consists only of hydrogen), the store of nuclear energy is

$$M_\odot \times 0.007c^2 \sim 2 \times 10^{30} \times 0.007\,(3 \times 10^8)^2 \text{ J}$$
$$\sim 1.3 \times 10^{45} \text{ J}.$$

The time of the active life of our "heavenly body" amounts to

$$\frac{M_\odot \times 0.007c^2}{J_\odot} \sim 3 \times 10^{18} \text{ s} \simeq 10^{11} \text{ years}.$$

It is more than enough to explain the observed effects.

9.3.2. Nuclear Reactions of the Proton-Proton Cycle. The scheme of this cycle is as follows:

$$p + p \to D + e^+ + \nu_e,$$
$$D + p \to {}^3\mathrm{He} + \gamma,$$
$${}^3\mathrm{He} + {}^3\mathrm{He} \to {}^4\mathrm{He} + 2p + \gamma.$$

The energy is liberated in the form of positrons e^+ (which are annihilated together with electrons to produce γ-quanta), γ-quanta, neutrino ν_e, and the kinetic energy of protons. The total energy liberated in this cycle is 28.5 MeV (this estimate was made in Sec. 9.3.1).

The probabilities of individual reactions of the cycle, or, which is the same, the mean durations of these reactions, are as follows:

$(p + p)$ — 10^{10} years,
$(D + p)$ — a few seconds,
$({}^3\mathrm{He} + {}^3\mathrm{He})$ — 10^6 years.

What are the conditions for a $p + p$ reaction to occur? These conditions were partially analyzed in Sec. 8.2. For the height of the repulsive Coulomb barrier in the $p + p$ process, the estimate $U_{pp} \sim 10^{-14}$ J was obtained. The temperature of the gas particles which can effectively surmount this barrier was estimated as $T \sim 10^9$ K.

However, the proton gas in the central region of the Sun has a temperature only of the order of 2×10^7 K, which sharply reduces the probability of the $p + p$ process. The probability of this process under these conditions, which corresponds to an under-the-barrier tunnel transition, was also estimated in Sec. 8.2. It was found to be $w_{pp} \sim 10^{-5}$. This means that one particle out of a hundred thousand particles is able to overcome the electrostatic Coulomb barrier.

We are interested not only in the process of surmounting the barrier, i.e. in approaching of two protons a distance at which strong interactions come into play. It is also necessary that due to a weak interaction, one of the protons undergo the following transformation[4]:

$$p \to n + e^+ + \nu_e,$$

i.e. be converted into a neutron since it is only the proton p and the neutron n that can form the nucleus of deuterium D. Such a combination of circumstances, viz. the passage through the barrier and the transformation of a proton into a neutron, is an event with an extremely low probability.

In spite of the fantastically low probability of the reaction $p + p \to D$, it ensures the required luminosity $J_\odot$ of the Sun. Let us prove this. The number of protons in the Sun has the following order of magnitude:

$$\frac{M_\odot}{m_p} \sim \frac{2 \times 10^{30} \text{ kg}}{2 \times 10^{-27} \text{ kg}} \sim 10^{57}.$$

The nuclear reaction $p + p \to D$ occurs on the average once in 10^{10} years. The number of individual acts of the

---

[4] A free proton is a stable particle. We are not considering a free proton but a proton participating in a strong interaction.

reactions $p + p \to D$ occurring in the Sun per second is

$$\frac{\text{Total number of protons}}{\left(\begin{matrix}\text{Number of nucleons}\\ \text{in a reaction cycle}\end{matrix}\right) \cdot \left(\begin{matrix}\text{Time of expectation of}\\ \text{a reaction in seconds}\end{matrix}\right)}$$

$$\sim \frac{10^{57}}{4 \times (10^{10} \times 3 \times 10^{7})} \sim 10^{39} \text{ acts/s.}$$

For the total emissive power of the Sun, we obtain

$$J_{\odot} \sim 28.5 \text{ MeV} \times 10^{39} \sim 28.5 \times 1.6 \times 10^{-13} \times 10^{39} \text{ W} \sim 3 \times 10^{27} \text{ W.}$$

This estimate is made under the assumption that the Sun is a purely hydrogen star. The measurements of the Sun luminosity give $J_{\odot} \simeq 4 \times 10^{26}$ J/s.

9.4. White Dwarfs

9.4.1. Possible Evolution of Stars of the Type of the Sun. At the centre of a normal star, hydrogen must burn out sometimes, and a helium core is formed, which is condensed to a large extent in the nuclear reactions. The *pp*-reaction will still occur in the outer layers of the star due to a reserve of hydrogen, and the envelope of the star expands inordinately, reaching a size comparable with that of the solar system. Such a formation is known as **red giant.**

As a red giant is cooled, its rarefied envelope dissipates. The remaining condensed core is called a **white dwarf.**

White dwarfs mainly consist of helium nuclei and free degenerate electrons. The electron mechanism of thermal conduction of white dwarfs ensures practically the same temperature over the entire volume of a star. No nuclear reactions occur in the bulk of such a star. It emits only due to the heat stored in it (the temperature drops from $T \sim 10^8$ K to $T \sim 0$ K). The time of cooling of a white dwarf is of the order of 10^8 years.

9.4.2. Density and Size of White Dwarfs. The density of stellar matter for which the electron gas has already become degenerate (the lower density limit) is determined from the relation

$$10^7 \rho^{5/3} \sim \frac{R}{M} RT, \qquad (9.4.1)$$

written in the SI units. Here, R is the gas constant, and M is the molar mass.

In order to calculate the left-hand side of (9.4.1), we must return to Sec. 5.10.6 where it was shown that the pressure of the degenerate quantum electron gas is

$$p_e \sim \varepsilon_F \left(\frac{N_e}{V}\right) \sim \frac{\hbar^2}{m_e} \left(\frac{N_e}{V}\right)^{5/3},$$

where m_e is the electron mass, N_e is the total number of electrons in a partially filled band of quantum states, and ε_F is the limiting energy of the electrons. Since we assume that the star consists of helium, one or two free electrons correspond to an atom, and the pressure of the matter is

$$p \sim \frac{\hbar^2}{m_e} \left(\frac{\rho}{m_N}\right)^{5/3} \sim 10^7 \rho^{5/3}, \tag{9.4.2}$$

where m_N is the mass of a nucleon, and ρ is the density of the matter. This relation follows from the obvious equalities

$$\frac{\rho}{m_N} = \frac{M_*}{V m_N} = \frac{N_N}{V} \approx \frac{N_e}{V}.$$

Let us return to Eq. (9.4.1). For $M = (1\text{-}4) \times 10^{-3}$, this relation gives

$$\rho \sim \left(\frac{RT}{10^4}\right)^{3/2}.$$

Putting $T \sim 10^7$ K in this expression, we obtain

$$\rho \sim \left(\frac{8.3 \times 10^7}{10^4}\right)^{3/2} \text{kg/m}^3 \sim 10^6 \text{ kg/m}^3.$$

At such a density of the hot matter, the electron gas has already become degenerate. For the Sun, the density is $\rho \lesssim 10^5$ kg/m^3 for $T_c \sim 2 \times 10^7$ K, while for white dwarfs, $10^7 < \rho < 10^9$ kg/m^3 at the same temperature.

It should be borne in mind that for an extremely degenerate electron gas ($T \ll T_d$), the temperature does not determine its state any longer. In particular, for $T = 0$ K, the equilibrium condition for a white dwarf is determined by the relation between the gravitational compressive pressure p_{grav} and the counterpressure p_{quant} which is not of

thermal nature but is determined by the quantum nature of the Fermi electron gas. Therefore,

$$p_{\text{quant}} \approx p_{\text{grav}},$$

or

$$10^7 \rho^{5/3} \sim G \frac{M_*^2}{R_*^4}. \tag{9.4.3}$$

Since $\rho \sim M_*/R_*^3$, this expression leads to

$$R_* \sim \frac{10^7}{GM_*^{1/3}}. \tag{9.4.4}$$

For white dwarfs, $M_* \approx M_\odot$, and the radius

$$R_* \sim \frac{10^7}{6.7 \times 10^{-11} (2 \times 10^{30})^{1/3}} \text{ m} \sim 10^4 \text{ km},$$

i.e. the size of white dwarfs is comparable with that of the Earth.

9.4.3. Limiting Mass of White Dwarfs. In a closely packed binary star consisting of a white dwarf and an ordinary star, the phenomenon of **accretion** can be observed. Under the effect of the strong gravitational field of a white dwarf,[5] matter flows from the ordinary star to the white dwarf. As the mass of the white dwarf increases, the density N_e/V of the electron gas grows, and hence its limiting momentum p_F increases too, which may become relativistic and even ultrarelativistic (see Sec. 5.10.6). This means that the limiting energy is

$$\varepsilon_F = cp_F = c\hbar \left(\frac{N_e}{V}\right)^{1/3} \gg m_e c^2.$$

The pressure of the ultrarelativistic electron gas is

$$p_e \sim \varepsilon_F \left(\frac{N_e}{V}\right) \sim c\hbar \left(\frac{N_e}{V}\right)^{4/3},$$

and the pressure of the matter is given by

$$p \sim c\hbar \left(\frac{\rho}{m_N}\right)^{4/3}.$$

---

[5] With a given mass of a star, the gravitational potential φ increases in its vicinity with decreasing size of the star ($|\varphi| = GM_*/R_*$).

The equilibrium condition of such superdense dwarf is written in the form

$$c\hbar \left(\frac{\rho}{m_N}\right)^{4/3} \approx G\frac{M_*^2}{R_*^4}.$$

Taking into account the expression $\rho \sim M_*/R_*^3$ for the density of the matter, we obtain

$$M_* \approx \left(\frac{c\hbar}{G}\right)^{3/2} \frac{1}{m_N^2}. \tag{9.4.5}$$

This is a remarkable formula expressing the mass of the white dwarf (whose electrons are in the limiting ultrarelativistic state) in terms of the universal constants and the mass of a nucleon. This limiting mass is $M_* = 1.4M_\odot$. The existence of the limiting mass for the white dwarf indicates that as it increases further, the pressure of the quantum ultrarelativistic electron gas cannot withstand gravitational forces any longer, and a rearrangement of the matter begins.

9.5. Superdense Neutron Stars

9.5.1. Size of Neutron Stars. At densities $\rho \sim 10^{12}$-10^{13} kg/m^3, the process of **neutronization** of matter begins, during which electrons are "thrust into" nuclei and transform protons into neutrons (emitting neutrinos). The excess of neutrons in nuclei leads to a reduction in the binding energy, and the nuclei decay, forming a neutron gas.

Neutrons, like electrons, are Fermi particles (have a half-integral spin), and the pressure of a degenerate neutron gas can be written by analogy with the formulas for electrons, i.e.

$$p_n \sim \frac{\hbar^2}{m_n}\left(\frac{N_n}{V}\right)^{5/3}.$$

Accordingly, the pressure of the matter will be

$$p \sim \frac{\hbar^2}{m_n}\left(\frac{\rho}{m_n}\right)^{5/3} = 10^4\rho^{5/3}.$$

This relation is written in the SI units. The radius of a neutron star can be determined by analogy with Eq. (9.4.4):

$$R_n \sim \frac{10^4}{GM_n^{1/3}}.$$

Thus, for neutron stars, we have $R_n \simeq 10$ km for $M_n \approx M_\odot$.

9.5.2. Rotation and Magnetic Fields of Neutron Stars. Suppose that a star with a mass of the order of that of the Sun loses stability after using up the nuclear sources of energy and is compressed under the effect of its own gravitation, becoming a neutron star. Since the angular momentum is conserved as a result of such a compression, according to Eq. (2.6.30), we have

$$M_\odot \omega_\odot R_\odot^2 = M_n \omega_n R_n^2, \tag{9.5.1}$$

where $M_\odot$ is the mass of the Sun (2×10^{30} kg), $\omega_\odot$ is the angular velocity of rotation of the Sun (3×10^{-6} rad/s), and $R_\odot$ is the radius of the Sun (7×10^8 m). Then for the angular velocity of the neutron star whose radius is $R_n \simeq 10$ km $= 10^4$ m, we obtain

$$\omega_n = \omega_\odot \left(\frac{R_\odot}{R_n}\right)^2, \quad \omega_n \sim \frac{3 \times 10^{-6} \times (7 \times 10^8)^2}{(10^4)^2} \text{ rad/s} \sim 10^4 \text{ rad/s}.$$

If we consider the hot plasma of the initial star as a perfectly conducting body (see Sec. 8.6), the intrinsic magnetic field of the star is as if frozen into its matter. Indeed, according to the law of electromagnetic induction, an emf is induced in a conductor crossing magnetic lines of force during its motion. Then in a perfect conductor (having an infinitely large conductivity), a very small emf would cause an infinitely high current, which is impossible. In other words, the magnetic lines of force of the star are as if glued to the particles of its plasma matter.

Therefore, the magnetic flux (see Eq. (3.3.1)) of a star cannot change as a result of its gravitational compression, i.e.

$$H_\odot R_\odot^2 = H_n R_n^2, \tag{9.5.2}$$

where $H_\odot$ is the magnetic field strength on the surface of the Sun ($\simeq 10^2$ A/m), and H_n is that of the neutron star. This gives

$$H_n = H_\odot \left(\frac{R_\odot}{R_n}\right)^2, \quad H_n \sim 10^2 \frac{(7 \times 10^8)^2}{(10^4)^2} \text{ A/m} \simeq 10^{12} \text{ A/m}.$$

It should be noted that the record values of magnetic field strength obtained in laboratory conditions in relatively

small volumes over a short time ($\sim 10^{-6}$ s) are of the order of 10^8 A/m.

9.5.3. Radio Emission by Pulsars. The direction of the magnetic moment $\mathbf{p}_m$ of a neutron star, as a rule, does not coincide with the direction of angular velocity ω. Since the star rotates, the vector $\mathbf{p}_m$ forms a conical surface in space. Under these conditions, the neutron star may become a source of the **magnetic dipole radiation**. As a matter of fact, the total electric charge of the star is zero to a high degree of accuracy, and the distribution of the matter in the bulk of the star is apparently spherically symmetric, which means that it has no electric dipole moment.

What parameters of a neutron star can determine the magnetic dipole radiation? This problem can be solved more easily in the CGS units. The frequency of such a radiation is obviously equal to the rotational frequency ω. The radiation intensity J must depend on the magnetic moment $\mathbf{p}_m$ and the rotational frequency. Besides, the expression for the intensity of electromagnetic radiation (including the magnetic dipole radiation) must contain the relativistic constant c. Consequently, J will depend on p_m, ω, and c.

Let us find this dependence by using dimensional analysis. For this purpose, we write

$$J \sim p_m^x \omega^y c^z,$$

where x, y, and z are certain unknown numbers. We write the dimensions of the quantities under consideration:

$$[J] = 1 \text{ erg/s} = 1 \text{ g} \cdot \text{cm}^2 \cdot \text{s}^{-3},$$

$$[p_m] = 1 \text{ erg/G} = 1 \text{ g}^{1/2} \cdot \text{cm}^{5/2} \cdot \text{s}^{-1} \text{ (see Appendix III)},$$

$$[\omega] = 1 \text{ s}^{-1}, \quad [c] = 1 \text{ cm} \cdot \text{s}^{-1}.$$

Substituting these dimensions into the initial expression, we obtain

$$1 \text{ g} \cdot \text{cm}^2 \cdot \text{s}^{-3} = 1 \text{ g}^{x/2} \cdot \text{cm}^{5(x+z)/2} \cdot \text{s}^{-x-y-z}.$$

Comparing the exponents of similar units, we obtain the following system of equations:

$$\left.\begin{aligned} \frac{x}{2} &= 1, \\ \frac{5(x+z)}{2} &= 2, \\ x+y+z &= 3. \end{aligned}\right\}$$

Solving this system, we get $x = 2$, $y = 4$, and $z = -3$. Therefore,

$$J \sim p_{\mathrm{m}}^2 \omega^4 c^{-3},$$

or

$$J \sim \frac{1}{c^3} p_{\mathrm{m}}^2 \omega^4. \tag{9.5.3}$$

This expression for the intensity of the magnetic dipole radiation can be transformed if we take into account the fact (see Eq. (3.5.4)) that the magnetic moment is

$$p_{\mathrm{m}} \sim H R_*^3. \tag{9.5.4}$$

Substituting this relation into (9.5.3), we obtain

$$J \sim \frac{1}{c^3} H^2 R_*^6 \omega^4. \tag{9.5.5}$$

This relation differs from the exact expression by a factor of 2/3.

Let us numerically estimate the total intensity of the magnetic dipole radiation of a neutron star. We shall use the familiar data: $H_n \sim 10^{10}$ Oe, $R_n \sim 10^6$ cm, and $\omega_n \sim 10^4$ s^{-1}. Substituting these values into Eq. (9.5.5), we obtain $J \sim 10^{40}$ erg/s $= 10^{33}$ W.

Is this value large? For the sake of comparison, we note that the total emissive power of all radio- and television transmitters on the Earth is of the order of a few thousand megawatts, or 10^9 W. In other words, the power of "cosmic radio transmitters" is fantastically high as compared with terrestrial sources.

The radiation of a neutron star is strictly directional (along the vector $\mathbf{p}_{\mathrm{m}}$). Therefore, we have the model of a "revolving beacon", and an observer on the Earth can de-

tect only separate pulses of radiation from such a star. For this reason, neutron stars are referred to as **pulsars.**

We can easily estimate the active lifetime of a pulsar. Indeed, the rotational energy is $E = I\omega_n^2/2$, where $I = M_n R_n^2$ is the moment of inertia of a sphere. Therefore,

$$E \sim M_n R_n^2 \omega_n^2 \sim 10^{30} \times (10^4)^2 \times (10^4)^2 \sim 10^{46} \text{ J}.$$

For the time τ of the rotational state of the pulsar, we obtain

$$\tau \sim \frac{E}{J} \sim \frac{10^{46}}{10^{33}} \sim 10^{13} \text{ s} \sim 10^6 \text{ years}.$$

During this time, the entire rotational energy of the pulsar will be transformed into the radiant energy.

9.5.4. Internal Structure of Neutron Stars. First of all, let us estimate the density of the matter constituting a pulsar. It is given by

$$\rho = \frac{M_n}{(4\pi R_n^3)/3} \simeq \frac{2 \times 10^{30}}{4\,(10^4)^3} \sim 5 \times 10^{17} \text{ kg/m}^3.$$

This is the density of a substance of atomic nuclei. For this reason, pulsars are sometimes referred to as "giant atomic nuclei".

In order to visualize the enormous value of this density of the matter of pulsars, let us estimate the weight of a cubic metre of this matter under terrestrial conditions. In the order of magnitude, its value is

$$\rho g \sim 10^{17} \times 10 \text{ N} \sim 10^{18} \text{ N} = 10^{14} \text{ tonf}.$$

We have already used the fact that the neutron matter of a pulsar is in a degenerate state. Let us prove this. According to Eq. (5.12.9), a system of particles obeys quantum laws if the quantum-mechanical wavelength of a particle is comparable with the mean distance between the particles, i.e.

$$\frac{\hbar}{p} \sim \left(\frac{N}{V}\right)^{-1/3}. \tag{9.5.6}$$

Here, p is the momentum of a neutron. Since the nucleons in a nucleus are relativistic particles, $p \sim m_n c$. The number of neutrons N/V per unit volume can be determined from the relation ρ/m_n.

Let us carry out numerical estimates: $\hbar/(m_n c) \sim 10^{-34}/(10^{-27} \times 10^8)$ m $\sim 10^{-15}$ m. The number density of neutrons in a pulsar is

$$\frac{N}{V} \sim \frac{\rho}{m_n} \sim \frac{10^{17}}{10^{-27}} \text{ neutrons/m}^3 \sim 10^{44} \text{ neutrons/m}^3.$$

The mean distance between the neutrons is

$$\left(\frac{N}{V}\right)^{-1/3} = \frac{1}{\sqrt[3]{N/V}} \sim \frac{1}{(10^{44})^{1/3}} \text{ m} \sim 10^{-15} \text{ m}.$$

Therefore, relation (9.5.6) is satisfied.

Since neutrons have a half-integral spin, the aggregate of neutrons should obey quantum Fermi statistics. Estimates show that correlated pairs of neutrons with antiparallel spins may appear in the neutron system of pulsars, these pairs being Bose particles (their total spin is zero). In other words, the neutron liquid may turn out to be superfluid.

The neutron matter of a pulsar contains an admixture of protons and electrons. Upon an appropriate paired correlation, protons may go over to the collective superconducting state. For electrons, such a state is impossible under the conditions in the pulsar.

The central regions of pulsars are apparently filled with a mixture of hyperons and mesons. In all probability, the outer layers of pulsars form a solid crust consisting of iron nuclei. The latter are packed most densely and hence have the maximum binding energy.

It is interesting to note that under the effect of ultrastrong magnetic fields of intensity $H \sim 10^{12}$-10^{14} A/m existing on the surface of pulsars, the matter of the crust considerably changes its properties. For example, the electron shells of iron atoms acquire an elongated needle shape in the direction of the field, and the matter of the crust forms a structure resembling that of polymers.

9.5.5. Gravitational Effects in the Vicinity of a Neutron Star. As regards gravitational effects, pulsars should be described by Einstein's relativistic theory of gravitation. Let us prove this by using very simple concepts. We shall introduce the concept of **weak** and **strong gravitational fields.** Proceeding from the well-known relations for the

motion of a body in a uniform gravitational field (see (Sec. 2.4)), we obtain

$$v_t = gt, \quad h_t = \frac{gt^2}{2}.$$

Eliminating time t from this system of equations, we get $v^2 = 2gh$. Substituting the relation $g \sim |\varphi|/h$, where φ is the gravitational potential,[6] into this expression, we obtain $v^2 \sim 2\,|\varphi|h/h \sim 2\,|\varphi|$. This relation can be written in the form

$$\frac{v^2}{c^2} \sim \frac{2|\varphi|}{c^2}. \tag{9.5.7}$$

It follows hence that when $v^2/c^2 \ll 1$, i.e. when the nonrelativistic theory of gravitation is applicable, the fields are considered to be weak (to be more precise, it is necessary that $|\varphi| \ll c^2$). For example, $|\varphi|/c^2 \sim 10^{-6}$ on the surface of the Sun, while $|\varphi|/c^2 \sim 10^{-3}$ on the surface of denser stars (white dwarfs).

Let us estimate the parameter $2\,|\varphi|/c^2$ for neutron stars. Using the expression for the potential φ of the gravitational field, we can write

$$\frac{2|\varphi|}{c^2} \sim 2\,\frac{G}{c^2}\,\frac{M_n}{R_n} \sim \frac{2\times 6.7\times 10^{-11}\times 2\times 10^{30}}{(3\times 10^8)^2\times 10^4} \simeq 0.3.$$

Thus, $|\varphi| \sim c^2$ for pulsars. Such gravitational fields are regarded as strong since external particles and bodies acquire $v \sim c$ in these fields (see Eq. (9.5.7)).

The obtained result indicates that the gravitational properties of pulsars should be described with the help of Einstein's relativistic theory of gravitation (also known as the **general theory of relativity**). According to this theory, the geometry of space in the vicinity of a neutron star must be non-Euclidean, i.e. *the space itself is considerably curved.* The pace of time in the vicinity of a pulsar will turn out to be slowed down.

---

[6] This expression is an analogue of the well-known relation $|E| = \Delta\varphi/\Delta x \sim \varphi/x$ between the strength E and the potential φ of an electrostatic field.

9.6. Gravitation and Relativity

9.6.1. Equivalence Principle. The existence of gravitational fields is a peculiar and universal phenomenon in nature. The properties of a gravitational field are established from the laws of motion of bodies in such fields.

According to the laws of mechanics, equations of motion for bodies in a weak gravitational field are written in the form

$$m_{\text{inert}}\mathbf{a} = m_{\text{grav}}\mathbf{g}, \tag{9.6.1}$$

where **a** is the acceleration, **g** is the gravitational field strength, m_{inert} is the **inertial mass**, i.e. the mass appearing, for example, in the relation $\mathbf{p} = m\mathbf{v}$ or $E = mc^2$, and m_{grav} is the **gravitational mass** playing the role of the gravitational charge producing a gravitational field (it appears, for example, in Newton's law $F = -Gmm_1/r^2$).

The basic property of a gravitational field is the uniformity of accelerations **a** of bodies in a field of a given strength **g** irrespective of their mass m. Taking Eq. (9.6.1) into consideration, we see that this property of the gravitational field leads to the following conclusion:

$$m_{\text{inert}} = m_{\text{grav}}, \tag{9.6.2}$$

i.e. *the inertial and gravitational masses are indistinguishable* (or *equivalent*).

This property of the gravitational field is unique. It should be noted, for example, that charges moving in an electrostatic field have an acceleration

$$\mathbf{a} = \frac{e}{m}\mathbf{E},$$

which depends on the charge-mass ratio.

At present, equality (9.6.2) has been experimentally verified to within 10^{-12}. This equality makes it possible to carry out an analogy between the analysis of the motion of bodies in inertial reference frames (IRF) in the presence of a gravitational field and the description of their motion in noninertial reference frames (NRF) in the absence of a gravitational field.

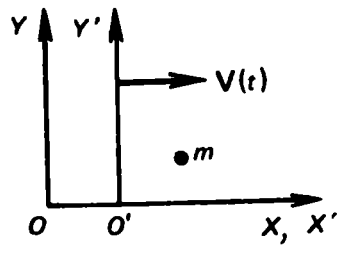


FIG. 9.1

Let us consider, for example, the motion of a uniformly accelerated NRF (Fig. 9.1). We introduce the following notation: $\mathbf{v}$ is the velocity of a particle of mass m in an IRF XOY, $\mathbf{v}'$ is the velocity of the same particle in an NRF $X'O'Y'$, and $\mathbf{V}(t)$ is the transport velocity of the NRF relative to the IRF, which uniformly varies with time. According to the classical law of velocity transformation, we have

$$\mathbf{v} = \mathbf{v}' + \mathbf{V}(t).$$

Accordingly, the equation of motion for the particle in the IRF has the form

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F}.$$

In the NRF, this equation can be written as

$$m\frac{d\mathbf{v}'}{dt} = \mathbf{F} - m\frac{d\mathbf{V}(t)}{dt}.$$

Thus, *in a uniformly accelerated NRF (when* $\mathbf{F} = 0$*), free bodies of any mass have the same constant acceleration* $-d\mathbf{V}(t)/dt$ *relative to this reference frame (which is equal and opposite to the acceleration of the reference frame itself).* The motion in a uniform constant gravitational field like the gravitational field of the Earth (over small regions where it can be treated as uniform) is of the same nature. Therefore, the motion relative to a uniformly accelerated NRF is equivalent to the motion in a uniform constant external gravitational field. Keeping in mind the general case as well, this statement can be referred to as the **equivalence principle.**

9.6.2. Geometry and Time in Noninertial Reference Frames. Let us consider an NRF $X'Y'Z'$ which uniformly rotates relative to an IRF XYZ about their common Z-axis (Fig. 9.2). We construct a circle in the $X'Y'$-plane of the NRF with the centre at the axis of rotation. In the absence of rotation, the ratio of the circumference L to its diameter D would be

π. However, during a rotation relative to the IRF, all the elements of length lying on the circumference experience the Lorentz contraction relative to the IRF, while the elements arranged along the diameter (at right angles to the velocity) remain unchanged. Consequently, L/D must differ from π. Therefore, *geometrical relations in NRF turn out to be non-Euclidean in contrast to those in IRF.*

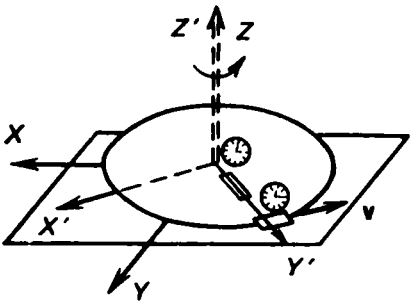


FIG. 9.2

If we consider two similar watches (rotating together with an NRF) so that one watch is on the circumference and the other at the centre, for an observer in an IRF, the former watch will be slowed down relative to the watch at the centre. The same situation should be observed in the NRF. Therefore, *the property of time changes as we go over to the NRF.*

In the above example, the NRF was a rotating reference frame. In such a system, a field of centrifugal forces emerges. These forces are zero at the centre of the circle and attain the maximum values on the circumference.

Applying the equivalence principle to this example, i.e. the statement that any NRF is equivalent to an IRF in the presence of a gravitational field, we arrive at the conclusion that *any gravitational field is just a change in the geometrical properties of space and time.*

9.6.3. Einstein's Equations. Since the sources of gravitational fields are the masses or energies of any kind of matter, the field equations themselves must determine the space-time metric in terms of the density of matter. Such equations were written for the first time in 1916 by Einstein. The structure of these equations is illustrated by the scheme shown in Fig. 9.3.

9.7. Expansion of the Universe

9.7.1. Friedman's Cosmological Solutions. Einstein's equations were solved in the framework of the homogeneous isotropic model of the Universe. The following approach may

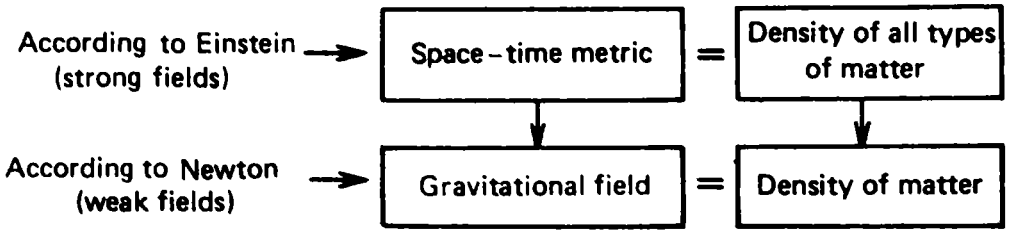


FIG. 9.3

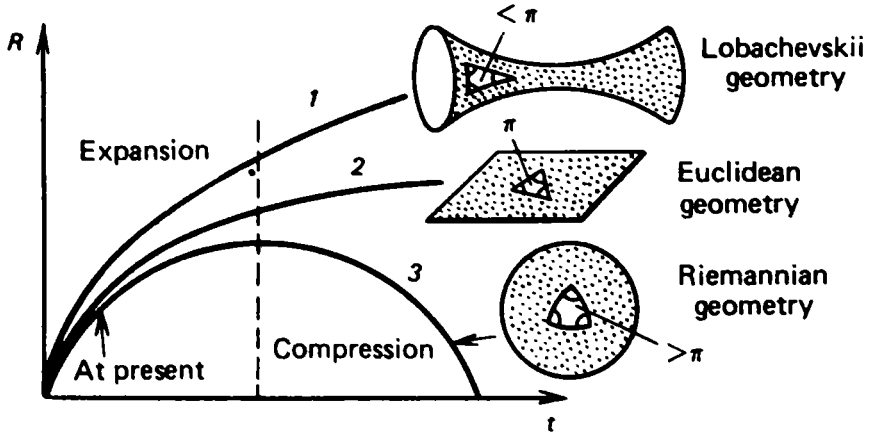


FIG. 9.4

serve as an analogy of these solutions. Let a body of mass m be in the gravitational field of a planet of mass M. Its initial velocity v may have different values. What is the dependence of the coordinates R of the body on time t? The energy conservation law gives the following relation:

$$m\frac{v^2}{2}-G\frac{mM}{R}=E,$$

or

$$\frac{m}{2}\left(\frac{dR(t)}{dt}\right)^2-G\frac{mM}{R(t)}=E.$$

The function $R(t)$ is the solution of this equation, its form depending on the values of the parameter E, namely (Fig. 9.4):

$$R(t)\rightarrow\begin{cases}\text{curve } 1 \text{ for } E>0,\\ \text{curve } 2 \text{ for } E=0,\\ \text{curve } 3 \text{ for } E<0.\end{cases}$$

The interpretation of these solutions as applied to the Universe will be given below. Near each curve in Fig. 9.4, there is a corresponding type of geometry for the Universe. The figure also shows the sum of the internal angles of a triangle for a given geometry.

9.7.2. Discovery of "Expansion" of the Universe. Astronomical observations led to **Hubble's empirical relation:**

$$v = HL, \tag{9.7.1}$$

where v is the velocity of recession of galaxies, L are intergalactic distances, and H is the coefficient known as the **Hubble constant.** It was established that

$$\frac{1}{H} \simeq 2 \times 10^{10} \text{ years.}$$

It is now convenient to interpret the solutions $R(t)$ as applied to the observed aggregate of galaxies forming the Universe. We shall simulate the actual three-dimensional space by a two-dimensional surface populated by two-dimensional creatures. Then curve *2* (see Fig. 9.4) corresponds to a plane unbounded surface for which all relations of Euclidean geometry are valid (in particular, the sum of the angles of a triangle is π). In this case, the distance L between point galaxies on the plane surface uniformly increases with time t. In order to clarify this, let us consider Fig. 9.5 in which the plane surface is represented by a rubber membrane. Let us imagine that this membrane is infinitely large and uniformly stretched in all directions. Then any of the numerous spots marked on the membrane can be chosen as the point of observation for a two-dimensional creature. Figure 9.5 shows that the farther any point galaxy from the observer, the higher the velocity of recession characterizing this point. We considered the two-dimensional model of an open Universe, i.e. the Universe in which the separation L between the galaxies grows indefinitely. Naturally, the radius R of the Universe also increases with time indefinitely.

Let us now consider a closed Universe. In Fig. 9.4, it corresponds to curve *3* for the function $R(t)$. The two-dimensional model of the closed three-dimensional space will now be the expanding surface of the sphere. In Fig. 9.6,

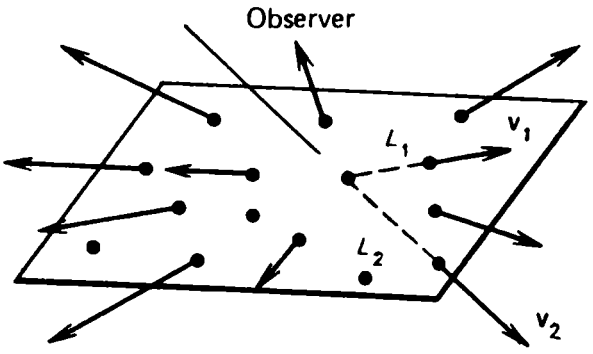


FIG. 9.5

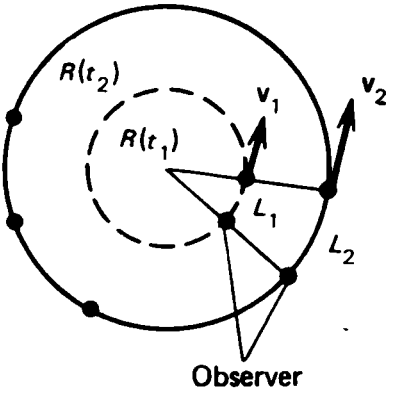


FIG. 9.6

we see a rubber balloon with point galaxies marked on its surface. The balloon expands, and its radius R (t) increases with time t. Here, the separation between points on its surface increases $(L_2 > L_1)$, as well as the velocities v of recession of point galaxies. The geometry on the surface of the sphere is non-Euclidean. For example, the sum of the angles of a triangle exceeds π. The dependence R (t) represented by curve *3* in Fig. 9.4 indicates that the expansion of the Universe attains a certain maximum value, and then it starts being compressed. The volume of such a Universe is bounded at each instant of time, as well as the amount of substance in it.

9.7.3. Critical Density. It follows from Hubble's law (9.7.1) that

$$v^2 = H^2L^2,$$

or

$$\rho v^2 = \rho H^2 L^2.$$

This is just the kinetic energy density of the galactic matter.

Let us denote by $\rho\varphi$ the density of energy of interaction between the galaxies, where φ is the gravitational potential. We assume that

$$\rho v^2 \sim \rho\varphi,$$

then

$$H^2L^2 \sim \varphi.$$

For the potential $\varphi = GM/L$, we can write the chain of equalities

$$\frac{\varphi}{L^2} = G\frac{M}{L^3} = G\rho.$$

Therefore, the previous relation will be written in the form

$$H^2 \sim \rho G. \tag{9.7.2}$$

This very important relation is a corollary of Hubble's law and the Friedman model of the Universe.

This model leads to a more exact relation between the averaged density ρ of the matter forming the Universe under observation and the parameters of motion H and interaction G of the matter. Namely, for a certain value of density ρ_{cr} which will be referred to as the **critical density**, the following relation is possible:

$$H^2 = \frac{8\pi}{3}\rho_{cr}G. \tag{9.7.3}$$

Hence for the critical density, we have

$$\rho_{cr} = \frac{H^2}{(8\pi/3)\,G},$$

$$\rho_{cr} = \frac{1}{(2\times 10^{10}\times 3\times 10^{7})^2 \times 8.37\times 6.67\times 10^{-11}}\ \text{kg/m}^3$$

$$= 0.5\times 10^{-26}\ \text{kg/m}^3.$$

The value of the critical density lies on the boundary separating the concepts of the open and closed models of the Universe. Since the parameters H and G are fixed, the measurement of the value of ρ will help to make a choice between these models.

The condition

$$H^2 > \frac{8\pi}{3}\rho G \tag{9.7.4}$$

corresponds to curve *1* in Fig. 9.4, i.e. in the language of semiqualitative concepts, we can say that the Universe must expand indefinitely if the energy of motion of the matter exceeds the energy of its interaction. The inverse inequality

$$H^2 < \frac{8\pi}{3}\rho G \tag{9.7.5}$$

corresponds to curve *3*, i.e. the closed model in which the expansion of the Universe should necessarily be replaced by its compression. This is the case when the energy associated with motion is lower than the energy of interaction.

What is known about the "smeared" density of the Universe, i.e. the density of observed matter averaged over the entire volume of the visible Universe? If we disregard the possible large mass of the neutrino and admissible in principle but hardly observed forms of matter like collapsed stars (black holes) and gravitational radiation, for the density we must put $\rho \sim 3 \times 10^{-28}$ kg/m^3. As a result,

$$H^2 - \frac{8\pi}{3}\rho G$$
$$= \left(\frac{1}{(2\times 10^{10}\times 3\times 10^7)^2} - 8.37\times 3\times 10^{-28}\times 6.67\times 10^{-11}\right) \text{ s}^{-2}$$
$$= (3\times 10^{-36} - 1.67\times 10^{-37}) \text{ s}^{-2} > 0,$$

i.e. the Universe only expands (of course, it is a preliminary conclusion).

9.8. Hot Universe

9.8.1. Discovery of Background Thermal Radiation. Since 1965, physicists and radioastronomers have been investigating the properties of **relict radiation.** It is an electromagnetic radiation lying in the radio-frequency range and corresponding to wavelengths $0.06 \text{ cm} < \lambda < 50 \text{ cm}$. The distribution of radiation intensity over wavelengths is of the type of Planck's distribution, and hence the radiation is in equilibrium and is associated with processes of thermal origin. The corresponding effective temperature of thermal radiation turned out to be 2.7 K. The radiation itself has the form of a background (noise) since it fills the entire space and is perfectly isotropic.

If we know the temperature of thermal radiation, we are in a position to determine all its characteristics. For this, we shall use formulas from Sec. 5.9. First of all, we shall determine the density of this radiation:

$$\frac{E}{V} = 4\frac{\sigma}{c}T^4$$
$$= \frac{4\times 5.67\times 10^{-8}}{3\times 10^8}(2.7)^4 \text{ J/m}^3 = 4\times 10^{-14} \text{ J/m}^3. \qquad (9.8.1)$$

The mean energy of a quantum is

$$\varepsilon = 2.82kT, \quad \varepsilon = 2.82 \times 1.38 \times 10^{-23} \times 2.7 \text{ J}$$
$$= 1.05 \times 10^{-22} \text{ J}, \tag{9.8.2}$$

the value of ε being determined by the position of the maximum on Planck's distribution curve $dE_\omega/d\omega$. Therefore, for the number of photons contained in a unit volume, we have

$$\frac{E/V}{\varepsilon} \simeq 4 \times 10^8 \text{ photons/m}^3. \tag{9.8.3}$$

Let us determine the numerical ratio of the density of matter and the density of the relict radiation of the Universe at the present time. In the previous section, we mentioned the density $\rho \sim 3 \times 10^{-28}$ kg/m^3 of matter averaged over the entire volume of the observed Universe. Then the number density of protons is

$$\frac{N_p}{V} = \frac{\rho}{m_p} = \frac{3 \times 10^{-28}}{1.67 \times 10^{-27}} \text{ protons/m}^3$$
$$\simeq 1.7 \times 10^{-1} \text{ protons/m}^3.$$

Therefore, for the ratio of the number densities of photons and protons, we obtain

$$\frac{n_\gamma}{n_p} \sim \frac{4 \times 10^8}{1.7 \times 10^{-1}} \sim 10^9.$$

9.8.2. Charge-Asymmetric Model of Early Universe. The obtained ratio n_γ/n_p is just the entropy of radiation per nucleon. Let us prove this by considering a plasma in which radiation dominates. We proceed from the general definition of temperature (5.5.1)

$$\frac{1}{T} = \frac{dS}{dE} \sim \frac{S}{E}.$$

Then the entropy is

$$S \sim \frac{E}{T}. \tag{9.8.4}$$

When applied to thermal radiation, this expression can be written in the form

$$\frac{S}{V} \sim \frac{E/V}{T}, \tag{9.8.5}$$

here, we go over to the entropy and energy of radiation per unit volume. Substituting (9.8.1) into (9.8.5), we obtain

$$\frac{S}{V} \approx 4 \frac{\sigma}{c} T^3. \tag{9.8.6}$$

One more circumstance is worth noting. The entropy in (9.8.4) is a dimensionless quantity if the temperature is expressed in energy units. If we want to have the dimensionless entropy and measure temperature in kelvins, we must divide both sides of (9.8.4) by the Boltzmann constant k:

$$\frac{S}{k} \sim \frac{E}{kT}. \tag{9.8.7}$$

Let us calculate the value of the dimensionless entropy of radiation per nucleon:

$$S_1 = \frac{S}{kn_p}, \tag{9.8.8}$$

where $n_p = N_p/V$ is the number density of nucleons, and $S \equiv S/V$. Substituting (9.8.6) into (9.8.8), we obtain

$$S_1 \approx \left(4 \frac{\sigma}{ck}\right) \frac{T^3}{n_p}. \tag{9.8.9}$$

Let us now return to the formula for the number of quanta per nucleon:

$$\frac{n_\gamma}{n_p} \equiv \frac{E/V}{\varepsilon n_p} = \frac{4\sigma T^4/c}{2.82kTn_p} = \frac{1}{2.82}\left(4 \frac{\sigma}{ck}\right) \frac{T^3}{n_p}. \tag{9.8.10}$$

In this formula, we have used expressions (9.8.1) and (9.8.2). Comparing (9.8.9) and (9.8.10), we finally get

$$\frac{n_\gamma}{n_p} \approx \frac{1}{2.82} S_1. \tag{9.8.11}$$

The exact expression is

$$\frac{n_\gamma}{n_p} = \frac{1}{3.7} S_1. \tag{9.8.12}$$

Let us now consider again the object under investigation, i.e. the Universe as a whole. Undoubtedly, the Universe expanded adiabatically, and in such a process the entropy is conserved. Therefore, the ratio n_γ/n_p becomes the most

important parameter characterizing cosmological expansion. Considering an early stage of the evolution of the Universe, we must write expression (9.8.12) in the form

$$S_1 = 3.7 \frac{n_\gamma}{\Delta n_N} \sim 10^9, \tag{9.8.13}$$

where Δn_N is the excess of baryons over antibaryons. At the early stage of expansion (for $T > 10^{13}$ K), about 10^9 nucleon-antinucleon pairs corresponded to one nucleon in the Universe at present. These pairs are annihilated (giving birth to n_γ), and the modern content of nucleons is characterized by an excess Δn_N of nucleons. This is the charge-asymmetric model of the early Universe.

A large body of experimental results obtained with the help of all facilities of astrophysical science and, in particular, by using satellite systems, space interplanetary apparatus, do not provide any proof of the existence of antimatter in space.

9.8.3. Change in Density and Temperature of Prestellar Matter in the Process of Cosmological Expansion. From the relation

$$H^2 = \frac{8\pi}{3} \rho G$$

and the time of expansion of the Universe

$$t = \frac{1}{H}$$

it follows that

$$\rho = \left(\frac{8\pi}{3} G t^2 \right)^{-1} \approx \frac{10^9}{t^2}. \tag{9.8.14}$$

It is the dependence of the mass density of matter (in kg/m^3) on time (in seconds) of cosmological expansion.

At the early stages of evolution of the Universe, the radiation mass density ρ_γ was considerably higher than the matter mass density ρ_{mat}. By the moment $t \sim 10^6$ years, the radiation and matter mass densities have become equal.

Since we are interested in the initial stages of the expansion, radiation predominated undoubtedly. According to

relativistic relations, the radiation mass density is defined as

$$\rho_\gamma = \frac{E/V}{c^2}.$$

Using formula (9.8.1) for the radiant energy density, we obtain

$$\rho_\gamma = 4\frac{\sigma}{c^3}T^4. \tag{9.8.15}$$

Finally, substituting (9.8.15) into the general relation (9.8.14), we obtain

$$T = \left(\frac{3c^3}{32\pi G\sigma}\right)^{1/4}\frac{1}{\sqrt{t}} \approx \frac{10^{10}}{\sqrt{t}}. \tag{9.8.16}$$

Here, the radiation temperature T is expressed in kelvins, and the time t in seconds.

Let us compile the table for the values of the density ρ and the temperature T of the prestellar matter depending on the time t elapsed after the beginning of expansion of the Universe (Table 9.1).

Table 9.1

t, s	ρ, kg/m^3	T, K
10^{-6}	10^{21}	10^{13}
10^{-4}	10^{17}	10^{12}
1	10^{9}	10^{10}
10^{6} years	10^{-18}	10^{3}
...	...	...

9.8.4. State of Aggregation at Early Stages of Evolution of Hot Universe. Let us consider the first row of numerical data of the table. Proceeding from the relation

$$m_N c^2 \sim kT,$$

we can estimate the temperature at which nucleons and antinucleons are born:

$$T \sim \frac{m_N c^2}{k}, \quad T \sim \frac{1.6\times 10^{-27}(3\times 10^8)^2}{1.4\times 10^{-23}}\ \text{K} \sim 10^{13}\ \text{K}.$$

The annihilation of nucleon-antinucleon pairs is accompanied by the radiation whose frequency is

$$\omega \sim \frac{m_N c^2}{\hbar},$$

$$\omega \sim \frac{1.6\times10^{-27}(3\times10^8)^2}{10^{-34}}\ \mathrm{s^{-1}} \sim 10^{24}\ \mathrm{s^{-1}}.$$

The corresponding wavelength is given by

$$\lambda = 2\pi\frac{c}{\omega},\quad \lambda \sim 6.3\,\frac{3\times10^8}{1.6\times10^{24}}\ \mathrm{m} \sim 10^{-15}\ \mathrm{m}.$$

In spite of a rapid expansion in the states corresponding to huge values of densities (see Table 9.1), nucleons, antinucleons, and radiation have time to attain thermodynamic equilibrium. Obviously, this system will also contain particles with a mass exceeding twice or thrice the mass of nucleons, i.e. the entire set of hyperons. Such a state of matter is realized in the time interval $10^{-6}\ \mathrm{s} < t < 10^{-4}\ \mathrm{s}$ from the beginning of the expansion. This time interval is known as **hadron era.**

Since unstable hyperons decay over a time of the order of 10^{-10} s, not only baryons (i.e. nucleons and hyperons) but also their decay products (K- and π-mesons) must exist in the hadron era. Mesons decay into leptons over a time of the order of 10^{-10}-10^{-8} s. Thus, the hadron era was formed by all kinds of particles: hadrons proper, leptons, and photons.

It should be noted that nucleonic nuclei cannot be formed in the hadron era since equilibrium photons with frequency $\omega \sim 10^{24}\ \mathrm{s^{-1}}$ have the energy

$$\hbar\omega \sim 10^{-34} \times 10^{24} \sim 10^{-10}\ \mathrm{J} \sim 10^9\ \mathrm{eV},$$

which is considerably higher than the binding energy of nuclei (for example, the binding energy of the deuterium nucleus consisting of a proton and a neutron is of the order of 2×10^6 eV). Such hard γ-quanta immediately destroy any nuclei.

Let us consider the next stage of the expansion (see the second and third rows of the table). During this period, the temperature varies in the interval $10^{10}\ \mathrm{K} < T < 10^{12}\ \mathrm{K}$. The energy is insufficient for generating hadrons. All the hyperons and mesons available have already decayed, and

the nucleon-antinucleon pairs have been annihilated. According to the charge-asymmetric model of the Universe, only a relatively small excess Δn_N of nucleons, i.e. protons and neutrons, is left. It should be recalled that the lifetime of free neutrons is of the order of 10^3 s.

The time interval 10^{-4} s $< t <$ 1 s from the beginning of the expansion is known as the **lepton era.** This name is justified since the overwhelming fraction of particles in this period is formed by stable leptons: electron-positron pairs, electron neutrinos, and antineutrinos. Photons, leptons, and the remaining unpaired nucleons are in thermal equilibrium.

By the end of the lepton era, neutrinos have practically stopped interacting with the matter of the system. Subsequently, the neutrino gas is just cooled (due to the adiabatic expansion). By the time of our era, the temperature of the neutrino gas must drop to $T \simeq 2$ K.[7]

Since the mean energies of thermal motion in the lepton era lie in the interval 10^6 eV $< kT < 10^8$ eV, the formation of nucleonic nuclei can be neglected.

In the next era of the expansion lasting over the time interval 1 s $< t < 10^6$ years, the radiation in equilibrium with a small admixture of electrons and nucleons mainly exists.

The time interval 1 s $< t <$ 100 s, during which the temperature lies in the interval 10^9 K $< T < 10^{10}$ K, is associated with an intense formation of the lightest nuclei: deuterium (a combination of a proton and a neutron) and helium (a combination of two protons and two neutrons). By the time $t \sim 10^3$ s, free neutrons, which had no time to take part in fusion reactions, have been transformed into protons, i.e. the nuclei of ordinary hydrogen. Calculations show that, as a result, the following proportion by nuclear mass is observed: 30% of helium nuclei and 70% of hydrogen nuclei. It should be noted that the results of analysis of the observed distribution of elements are in good agreement with the predictions of the hot model of the Universe.

During the next 10^6 years, the temperature of the prestellar matter still remains rather high (it varies within the

---

[7] The detection of such "cold" neutrinos is in all probability a problem to be solved by experimenters of future.

interval 10^3 K $< T < 10^9$ K). The state of aggregation of this medium is the plasma. The recombination process in which protons, deuterons, and helium nuclei are combined with free electrons to form neutral atoms of hydrogen and helium starts only at temperatures 3500-4000 K. Equilibrium photons with a temperature $T < 3500$ K are now unable to ionize hydrogen and helium atoms. The photon gas practically stops interacting with the matter and forms the background of thermal radiation which, upon cooling during the subsequent adiabatic expansion, has acquired the temperature $T \simeq 2.7$ K by our time. For this reason, the detected background thermal radiation is termed the relict radiation, i.e. the one left from ancient times.

9.9. Fusion of Elements in Stars

At a certain stage of the general cosmological expansion (at any rate, after the recombination epoch), not only density fluctuations but also collective hydrodynamic processes must become effective in the prestellar gaseous medium. Under the conditions when gravitation plays a significant role, these phenomena change their nature. For example, a sufficiently large density fluctuation may result in the formation of a stable cluster of a gaseous substance. The formation of galaxies and their clusters forming the large-scale cellular structure of the Universe is thoroughly studied by researchers. Many aspects of this problem still remain unclear.

As regards the formation of stars, a much better understanding is attained in this more concrete problem. However, the progress made by theoretical physicists in this field is hampered because of a lack of experimental data. As a matter of fact, stars in the process of generation (which proceeds via gravitational condensation of gases) are still quite "cold" and emit infrared radiation. This radiation is absorbed by the Earth's crust, and hence detectors sensitive to infrared radiation should be taken out of the limits of the atmosphere.

After these remarks concerning the formation of galaxies and stars, let us consider the fusion of nuclei of elements in stars. It was mentioned above that only the fusion of deuterium and helium nuclei could take place at the pre-

stellar stage. Therefore, the "building material" for the stars of the first generation was hydrogen (normal and heavy) as well as helium.

In the stars of the type of the Sun, the proton-proton cycle considered above led to the formation of helium from hydrogen. Let us recall the scheme of reactions of the *pp*-cycle, which may occur even in the purely hydrogen medium:

$$p + p \rightarrow D + e^+ + \nu,$$
$$D + p \rightarrow {}^3\mathrm{He} + \gamma,$$
$${}^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + 2p.$$

This *pp*-cycle is an "additional caterer" of helium nuclei and introduces a small correction to the data on the abundance of elements in the Universe.

In ordinary stars which, however, differ from the Sun in a higher number density of ^{4}He and have higher temperatures of central regions, the following reaction becomes possible:

$${}^3\mathrm{He} + {}^4\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma.$$

The unstable beryllium isotope ^{7}Be may either capture an electron to form the stable lithium isotope ^{7}Li

$${}^7\mathrm{Be} + e^- \rightarrow {}^7\mathrm{Li} + \nu$$

or (at a higher temperature) combine with a proton to form the radioactive boron isotope ^{8}B:

$${}^7\mathrm{Be} + p \rightarrow {}^8\mathrm{B} + \gamma.$$

The latter decays according to the scheme

$${}^8\mathrm{B} \rightarrow {}^8\mathrm{Be} + e^+ + \nu.$$

This decay is remarkable since it is accompanied by the emission of high-energy neutrinos (on the average, with an energy of 8-9 MeV) which can be detected on the Earth.

In the stars where hydrogen has been burnt out and the core consists of two-thirds of helium, the nature of evolution drastically changes: the core becomes dense, and its temperature increases to $T \sim 10^8$ K, while the outer en-

velope expands. This is the red giant, and the following reaction is possible in its core:

$$^4\text{He} + ^4\text{He} \rightarrow ^8\text{Be} + \gamma.$$

However, the ^{8}Be nucleus is highly unstable and decays over a time of the order of 10^{-15} s. Nevertheless, the ^{8}Be nucleus could manage to capture another ^{4}He nucleus to form ^{12}C:

$$^8\text{Be} + ^4\text{He} \rightarrow ^{12}\text{C} + \gamma.$$

For this reaction to be rapid, the ^{12}C nucleus must have a resonance energy level (i.e. the level very close to the energy of the initial system of nuclei) corresponding to 7.82 MeV. The existence of such an energy level for the ^{12}C nucleus was significant for theoretical scientists since, otherwise, the chain of formation of all other elements of the Periodic System would terminate. And indeed, this energy level of the ^{12}C nucleus was identified by experimenters under laboratory conditions.

This is a peculiar circumstance. It turns out that the nature of our Universe considerably depends on the exact value of a single energy level of a single nucleus.

Now, the following nuclear reaction can take place:

$$^{12}\text{C} + ^4\text{He} \rightarrow ^{16}\text{O} + \gamma.$$

The excited energy level of the ^{16}O nucleus, corresponding to this reaction, is slightly lower than the resonance level, and hence the reaction rate is not very high. This fact explains nearly the same abundance of carbon and oxygen in nature and, in particular, makes the life in the form known to us possible.

It should be noted that ^{12}C and ^{16}O nuclei, like the ^{4}He nucleus, are quite stable, i.e. have a high binding energy per nucleon. Moreover, the nuclei $^4_2$He$(2p, 2n)$ and $^{16}_8$O$(8p, 8n)$ belong to the category of **double-magic nuclei** (their proton and neutron shells are completely filled).

The fusion of the next heavier nucleus involving helium nuclei requires still higher energies of the nuclei participating in the fusion reaction. This is natural since as the atomic number of an element increases, the charge of the nucleus increases together with the energy barrier that should be surmounted by a helium nucleus for its fusion with heavier nucleus.

Therefore, the temperatures $T \sim 10^9$ K are required for the initiation of combustion reactions of ^{12}C and ^{16}O. Such temperatures are attained in giant stars in which these reactions may proceed via numerous channels:

$$^{12}C + {}^{12}C \rightarrow \begin{cases} {}^{24}Mg + \gamma, \\ {}^{23}Na + p, \\ {}^{20}Ne + {}^{4}He; \end{cases} \qquad {}^{16}O + {}^{16}O \rightarrow \begin{cases} {}^{32}S + \gamma, \\ {}^{31}P + p, \\ {}^{31}S + n, \\ {}^{28}Si + {}^{4}He. \end{cases}$$

The main product of combustion reactions of carbon and oxygen is ^{28}Si. The binding energy per nucleon in the ^{28}Si nucleus is especially high.

Going over to higher temperatures $T > 3 \times 10^9$ K, we observe combustion reactions of silicon:

$$^{28}Si + {}^{4}He \rightleftarrows {}^{32}S + \gamma,$$

$$^{32}S + {}^{4}He \rightleftarrows {}^{36}Ar + \gamma,$$

and so on. It should be noted that the energy of γ-quanta emitted in these reactions by the excited nuclei is sufficiently high to split these nuclei.

Ultimately, the reactions of this type lead to the formation of ^{56}Fe and elements close to it. These elements are characterized by the maximum binding energy per nucleon. For this reason, fusion reactions accompanied by energy liberation terminate at ^{56}Fe.

The fusion of elements heavier than iron involves a different mechanism. This mechanism is associated with a slow capture of neutrons by nuclei.[8] In such reactions, free neutrons are captured by nuclei infrequently. Therefore, if unstable nuclei are formed, they have time to decay before they capture the next neutron.

The results of observations confirmed that the fusion reactions involving a slow capture of neutrons can be realized in stars. For example, technetium does not have stable isotopes and is not encountered in the natural state on the Earth (the half-life of $^{99}_{43}Tc$ is 2×10^5 years). Technetium ^{99}Tc is formed during a slow capture of neutrons and was

[8] Being a neutral particle, a neutron does not encounter a repulsive Coulomb potential barrier while penetrating a nucleus.

detected in red giants. The spectra of red giants also contain promethium $^{145}_{61}Pm$ (its half-life is 18 years). In general, it should be noted that red giants of various types are real "plants for producing elements".

After these remarks, a natural question arises: what is the source of free neutrons in stars? If we consider relatively high temperatures $T \geqslant 10^9$ K, the following reactions can be the sources of neutrons:

$$^{13}C + {}^4He \rightarrow {}^{16}O + n,$$

$$^{21}Ne + {}^4He \rightarrow {}^{24}Mg + n,$$

and so on. At lower temperatures, the following reactions may occur in ordinary fixed stars:

$$^{13}C + \gamma \rightarrow {}^{12}C + n,$$

$$^{14}N + \gamma \rightarrow {}^{13}N + n.$$

These high-temperature reactions in giant stars produce neutron fluxes with a density $\leqslant 10^{20}$ neutrons/(m^2·s). For comparison, we note that in laboratory-scale reactors on the Earth (in their active regions), neutron flux densities of 10^{17}-10^{19} neutrons/(m^2·s) have been attained, and in pulsed reactors, this value is of the order of 10^{21} neutrons/(m^2·s).

The following scheme can serve as an example of a consecutive slow capture of neutrons by nuclei:

$$^{56}_{26}Fe + n \rightarrow {}^{57}_{26}Fe + \gamma,$$

$$^{57}_{26}Fe + n \rightarrow {}^{58}_{26}Fe + \gamma,$$

$$^{58}_{26}Fe + n \rightarrow {}^{59}_{26}Fe + \gamma,$$

$$^{59}_{26}Fe \rightarrow {}^{59}_{27}Co + e^- + \tilde{\nu}.$$

Here, the unstable ^{59}Fe nucleus undergoes a β^--decay, i.e. the decay in which a neutron of the nucleus undergoes the following transformation:

$$n \rightarrow p + e^- + \tilde{\nu}.$$

The transformation chain of the ^{56}Fe nuclei into heavier elements terminates at lead. In this process, the multiple slow capture of neutrons alternates in this chain with acts of a β^--decay. The reason behind this termination of the chain consists in that the heaviest isotope obtained as a re-

sult of a slow capture of neutrons by nuclei is ^{209}Bi. The ^{209}Bi nucleus combined with a neutron undergoes the following transformations:

$$^{209}Bi + n \rightarrow {}^{210}Bi,$$

$$^{210}Bi \rightarrow {}^{4}He + {}^{206}Pb.$$

During a capture of neutrons, nuclei with masses larger than that of ^{209}Bi undergo not a β^--decay but a decay involving helium nuclei. This completes the fusion of elements in the process of evolution of fixed stars.

Then how were the heaviest elements formed? If the density of free neutrons in stellar matter increases, the conditions can be attained under which the nucleus of an element has time to capture consecutively many neutrons before it decays (if the nucleus is unstable). For these **fast neutron capture reactions** to occur, neutron flux densities of 10^{31}-10^{44} neutrons/($m^2 \cdot s$) are required. It should be noted that the neutron flux density during artificial nuclear explosions on the Earth attains values of 10^{35} neutrons/($m^2 \cdot s$). Thus, fast neutron capture reactions are impossible even in the massive fixed stars. As was mentioned above, neutron flux densities in such stars are of the order of 10^{20} neutrons/($m^2 \cdot s$).

Huge bursts of supernovas are the regions in the Universe where the neutron fluxes have the required values of densities for the chain of consecutive acts of a fast capture of neutrons by heavy nuclei to take place. According to modern ideas, supernovas are the stars that have experienced a long evolution and contain a considerable fraction of heavy elements. The supernovas spectra have broad absorption lines of ionized Fe, Ca, Si, and neutral ^4He atoms but do not contain bright lines corresponding to H. The emissive power during the burst of a supernova exceeds the emissive power of the Sun by a factor of 10^9. The velocity of motion of the gaseous envelope stripped by the explosion attains the values of 10-20 000 km/s. The glowing sphere (shock front) at the maximum of luminosity attains the size exceeding the size of the Earth's orbit by a factor of ten.

The shock wave emerging at the surface of an exploding star causes the energy cumulation in relatively thin layers of stellar matter (see Sec. 6.11.5). High-power neutron fluxes

generated in this case may ensure the intense processes of a fast capture of neutrons by nuclei leading to the formation of the heaviest elements.

It should be noted that a fast capture of many neutrons by heavy nuclei was realized in terrestrial conditions during underground nuclear explosions. Under the effect of extreme fluxes of neutrons, a chain of transuranic elements up to fermium was obtained. The process was as follows: natural uranium captured neutrons generated during the explosion

$$^{238}_{92}\mathrm{U} + 17n \rightarrow {}^{255}_{92}\mathrm{U},$$

forming an overloaded nucleus, followed by a cascade of eight β^--decays:

$$^{255}_{92}\mathrm{U} \xrightarrow{\beta^-} {}^{255}_{93}\mathrm{Np} \xrightarrow{\beta^-} {}^{255}_{94}\mathrm{Pu} \xrightarrow{\beta^-} {}^{255}_{95}\mathrm{Am} \xrightarrow{\beta^-} {}^{255}_{96}\mathrm{Cm} \xrightarrow{\beta^-}$$

$$^{255}_{97}\mathrm{Bk} \xrightarrow{\beta^-} {}^{255}_{98}\mathrm{Cf} \xrightarrow{\beta^-} {}^{255}_{99}\mathrm{Es} \xrightarrow{\beta^-} {}^{255}_{100}\mathrm{Fm}.$$

CONCLUDING REMARKS

We have now come to the end of our trip to the land of physics. This is by no means an excursion in which the author as a guide conducts the students as tourists, since you have to work yourself if you want to see something here.

On this short trip, the author has endeavoured to acquaint the reader with many physical phenomena. All these phenomena had to be understood and assigned appropriate positions in the general physical pattern of nature. This has become possible owing to the fact that the present level of advancement of physics allows a unique interpretation of quite diverse phenomena and processes.

Let us outline the main trends in the development of natural sciences in general and physics in particular. These include elementary particle physics, astrophysics and cosmology, and molecular biology. A cursory glance shows that these disciplines cover outwardly contradictory topics, viz. the study of the smallest, the largest, and the most complicated states of matter. However, the entire history of the evolution of science shows that the key to the solution of the most intricate problems lies in the dialectical union of opposites. Hence the role of the dialectical materialistic views, which form the basis of scientific knowledge, becomes quite significant. We are on the verge of not only a technological era heralding the scientific and technical progress in industry and agriculture but also an era of outstanding discoveries throwing light on the hidden treasures of knowledge.

APPENDICES

I. Fundamental Experiments in the History of Physics and Outstanding Inventions in Physics and Engineering

Paradoxical though it may seem, there is an element of subjectivism in the choice of fundamental experiments in contrast to the list of theoretical achievements generalizing experimental results.

Mechanics

1. Galileo's experiments on motion of bodies (Italy, 1638).

In Galileo's opinion, a well-staged experiment and its appropriate mathematical analysis may be decisive for understanding a physical phenomenon.

2. The discovery of a planet (later called Neptune) whose existence was first predicted theoretically (Le Verrier, France, and Galle, Germany, 1846).

Calculations made by Le Verrier on the basis of Newtonian mechanics and the law of universal gravitation were verified in observations by Galle.

Relativistic Principles

1. Michelson's experiments and the discovery of the independence of the velocity of light from the motion of a source (USA, 1887).

Our planet was used as a moving source of light. Optical measurements revealed the invariance of the velocity of light, which was found to be the limiting velocity.

2. The discovery of the predicted mass defect of atomic nuclei (Aston, England, 1919).

It was one of the numerous verifications of Einstein's relativistic mechanics. Mass spectrography founded by Aston played an outstanding role in the development of nuclear physics.

3. The invention of the first circular-orbit accelerator of elementary particles, viz. cyclotron (Lawrence, USA, 1932).

Cyclotrons, which are relativistic instruments, remain the basic tools for studies in elementary particle physics.

Electromagnetism and Optics

1. Faraday's experiments on electromagnetism (England, 1830).

Faraday was born in a poor family and worked as an apprentice to a book-binder. Self-education and perseverance made him an outstanding experimenter. He was also a leading popularizer of science.

2. Hertz's experiments on the detection of predicted electromagnetic radiation (Germany, 1888).

The existence of the radiation field followed from Maxwell's electrodynamics. Hertz's experiments proved that the field does exist as a physical reality.

3. The invention of optical microscope (Leeuwenhoek, Holland, 1673).

Leeuwenhoek was the first to observe the new world, viz. the world of microscopic organisms.

4. The creation of electron microscope (Zworykin, USA, 1942).

The inventor was born and educated in Russia. In 1917, he emigrated to the USA.

5. The invention of optical telescope (Galileo, Italy, 1608).

This invention paved the way to the investigation of the new world of celestial bodies. Galileo was the first to observe lunar craters and satellites of Jupiter.

6. Radio telescope (Jansky, USA, 1931).

The "image" of the Universe formed by radio waves was an epoch-making achievement in astronomy. Quasars, pulsars, relict radiation, and other objects were discovered by radioastronomical methods.

7. The creation of basic devices in electrical engineering (Dolivo-Dobrovolskii, Russia, 1891).

This scientist developed the three-phase current system, constructed the first three-phase a.c. generator with rotating magnetic field, and developed all the elements of three-phase circuits.

8. The development of radio engineering (Popov, Russia, 1895).

The very first application of radiocommunication facilities helped save the lives of fishermen stranded on an ice floe.

9. Television (Zworykin, USA, 1931).

As was mentioned above, the scientist was born in Russia.

10. Holography (Gabor, England, 1948, and Denisyuk, USSR, 1962).

Gabor was born in Hungary. He laid the theoretical foundation of holographic image formation. These ideas could be realized in practice only after the creation of lasers, viz. the sources of coherent radiation. The Soviet physicist Denisyuk proposed the holographic method of recording in three-dimensional media, which made it possible to record the amplitude, phase, and the spectral composition of the wave field of an object and to obtain its undistorted space image. The holograms obtained by this method can be reconstructed by using an ordinary source of radiation with a continuous spectrum.

Gravitation

1. Cavendish's experiments on the measurement of gravitational constant (England, 1798).

The measurements carried out by Cavendish in his laboratory allowed our planet to be "weighed" for the first time.

2. Experimental verification of the equality of inertial and gravitational masses (Eötvös, Hungary, 1896, and Braginskii, USSR, 1971).

The accuracy of measurements made by Eötvös was 10^{-9}, while in Braginskii's measurements it was of the order of 10^{-12}. The fact of the equality of inertial and gravitational masses forms the basis of Einstein's general theory of relativity, viz. the relativistic theory of gravitation.

3. Experiments on measuring the predicted change in the photon frequency in the gravitational field (Pound and Rebka, USA, 1960).

The measurement of the photon frequency was carried out in a 20-m high vertical tower. The accuracy of measurements was about 10^{-15}. This fantastic accuracy could be attained owing to the effect of resonance absorption of gamma radiation by the nuclei of the lattice (Mössbauer effect). The dependence of the photon frequency on the gravitational field intensity follows from Einstein's theory of gravitation.

4. The observation of the predicted expansion of the Universe (Hubble, USA, 1929).

Scientific foundations for the analysis of the large-scale structure of the Universe were formed in the framework of the relativistic theory of gravitation. The Russian physicist and mechanical engineer Friedman applied the equations of the general theory of relativity to the analysis of the Universe as a whole. In the framework of the homogeneous and isotropic model of the Universe, he discovered its instability (USSR, 1922).

5. The observation of the predicted background relict radiation (Penzias and Wilson, USA, 1965).

This radiation, which is a residue of the early epoch of compressed hot Universe, was predicted in 1946 by Gamow, the Russian theoretical physicist who did a lot of work in his motherland (the discovery of the quantum-mechanical tunnel effect in 1928 and the explanation of the radioactivity of nuclei). Gamow was not only the author of the generally accepted model of the hot Universe (1946) but also the first to decipher the genetic code of the DNA molecule (1954). In 1934, he emigrated to the USA where he died in 1968.

6. The launching of the first Earth's satellite (Korolev and coworkers, USSR, 1957).

This was the beginning of the space era. The next stages were the manned flight to space (Gagarin, USSR, 1961) and the expedition to the Moon (Armstrong's crew, USA, 1969).

Quantum Mechanics

1. Rutherford's experiments on the investigation of the atomic structure. The discovery of atomic nucleus (England, 1911).

Rutherford proposed the planetary model of the atom. This model was theoretically analyzed by Bohr (1913-1923) who laid the foundation of the quantum concepts in physics.

2. The discovery of spectral series of the hydrogen atom (Balmer, Switzerland, 1885, Lyman, Germany, 1906, and Paschen, Germany, 1908).

The discrete nature of the emission spectra of atoms indicated that the values of energy in an atom are quantized. The analysis of

spectra was not only the starting point of the theory but also the touchstone for theoretical constructions.

3. Franck's and Hertz's experiments on measuring the excitation potentials in gases (Germany, 1914).

This was another experimental method proving the discreteness of the energy values in an atom.

4. The Stern and Gerlach experiments on the space quantization of atomic moments (Germany, 1922).

The most important conclusion following from these experiments was the one concerning the presence of spin of particles in addition to their mass and charge.

5. Experiments on electron diffraction by crystals (Davisson and Germer, USA, 1927).

It is interesting that these experiments were carried out after the mathematical apparatus of quantum mechanics had been developed by Heisenberg (Germany, 1925) and Schrödinger (Austria, 1926).

6. The invention of maser, the quantum generator of radio waves (Prokhorov and Basov, USSR, 1954, and Townes, USA, 1954).

7. The creation of laser (Maiman, USA, 1960).

Quantum Electrodynamics

1. Compton's experiments on photon scattering by electrons (USA, 1923).

These experiments convincingly proved that a quantum of electromagnetic radiation behaves as a particle.

2. The discovery of the predicted antielectron (positron) and of annihilation (Anderson, USA, 1932).

The synthesis of the theory of relativity and the quantum theory made it possible for the English theoretical physicist Dirac to proclaim the existence of the world of antiparticles (1928-1930).

3. Lamb's and Rutherford's experiments on the fine structure of energy levels in the hydrogen atom (USA, 1947).

These experiments are remarkable in two respects. Firstly, according to Dirac, the energies corresponding to the states $2S_{1/2}$ ($n = 2$, $l = 0$, $j = 1/2$) and $2P_{1/2}$ ($n = 2$, $l = 1$, $j = 1/2$) in the hydrogen atom must have the same value in contrast to the results of the experiments. This contradiction was removed by introducing the concept of physical vacuum affecting the properties of atomic electrons. Henceforth, this trend of the theory led to the creation of modern quantum electrodynamics (Feynman, Schwinger, and Tomonaga, USA and Japan, 1949). Secondly, the Lamb-Rutherford experiments themselves have become possible owing to the development of super-high-frequency technology and the creation of radiospectroscopy methods.

Elementary Particle Physics and Nuclear Physics. Extreme States of Matter

1. The discovery of the neutron (Chadwick, England, 1932).

This discovery was decisive for the understanding of nuclear structure.

2. The discovery of the predicted neutrino (Reines and Cowen, USA, 1956).

The analysis of the β-decay of nuclei revealed the violation of the energy conservation law. In 1931, the Swiss theoretical physicist Pauli introduced a new particle, viz. neutrino, to save this law. The development of high-power neutrino sources (nuclear reactors) has made their detection possible.

3. The violation of parity in weak interactions (Wu, USA, 1957).

This research work was proposed by Lee and Yang (theoretical physicists of Chinese origin working in the USA). The experiments forced the leading physicists in many countries (for example, Landau in the USSR) to review the theory and refute a number of generally accepted concepts.

4. Experimental detection of massive intermediate vector bosons (Rubia who headed the large international group of physicists in Geneva, CERN, 1983).

This research was carried out on a giant proton-antiproton colliding beam accelerator constructed with the collaboration of all West-European countries. The energy of a $p\tilde{p}$-collision is about 500 GeV, or 5×10^{11} eV. The mass of intermediate bosons was found to be hundred times the mass of a proton.

The experimental detection of intermediate bosons which are responsible for weak interactions confirmed once again the prediction concerning the unified theory of electroweak interactions developed by Weinberg and Glashow from the USA and Salam from Pakistan. For energies of the order of 100 GeV, quantum electrodynamics requires a refinement: two charged particles exchange not only a photon but also a neutral intermediate boson (at low energies, the latter mechanism is strongly suppressed). As a result, at distances of the order of 10^{-16} cm, weak and electromagnetic interactions become practically similar in intensity and can be described uniquely. The essence of this description lies in that the interactions of particles at such short distances are highly symmetric: there exist four equivalent massless fields which are transformed into three massive intermediate bosons W^+, W^-, and Z^0 and a massless γ-photon at large distances in view of the **spontaneous violation of symmetry**. Under normal conditions (large scale), intermediate bosons exist only virtually and cannot be observed. In this case, the difference between weak and electromagnetic interactions is enormous. The new fundamental theory not only embraces all known phenomena in this field but also predicts new types of phenomena.

5. Discovery of predicted neutron stars (Hewish and Bell, England, 1967).

The theoretical concepts concerning neutron stars were developed by Landau (USSR, 1932). The advances in radioastronomical research permitted periodically pulsating radiation sources to be registered. Further observations and their analysis revealed that these sources are rapidly rotating magnetized neutron stars.

6. The creation of nuclear reactor (Fermi, USA, 1942, and Kurchatov, USSR, 1946).

Nuclear reactors have become a practically inexhaustible source of energy.

7. Instruments based on the Mössbauer effect (Germany, 1958).

The potentialities of these instruments were outlined in the book.

Macrophysics. Thermal Phenomena

1. Perrin's experiments on the distribution of suspended particles in a liquid. The determination of atomic mass and Avogadro's number (France, 1908).

These experiments were stimulated by Einstein's theoretical work on Brownian movement (1905). The research work carried out by Einstein and Perrin provided a convincing proof of the atomic structure of matter and of the fact that thermal phenomena boil down to random motion of all the particles constituting a body.

2. The discovery of superfluidity of liquid helium (P. Kapitza, USSR, 1938).

This discovery was of fundamental importance for quantum macrophysics. Its theoretical analysis was made by Landau (USSR, 1941), Bogoliubov (USSR, 1947), and Feynman (USA, 1957).

3. The invention of the bubble chamber (Glaser, USA, 1952).

This instrument continues to be the main detector of particles in high-energy physics.

4. Steam piston engine (Watt, England, 1784, and Polzunov, Russia, 1765).

5. Steam turbine (Parsons, England, 1884, and Laval, Sweden, 1893).

6. The invention of the internal combustion engine (Otto, Germany, 1876, and Kostovich, Russia, 1880).

7. The creation of air-jet engines (Lyul'ka, USSR, 1939, Whittle, England, 1941, and Franz, Germany, 1942).

8. The creation of liquid-propellant rocket engines (Goddard, USA, 1926, Oberth, Germany, 1929, and Glushko, USSR, 1931).

Solid State Physics

1. The development of X-ray diffraction analysis (Laue, Germany, 1912).

This experimental method has made it possible to "see" that a crystal is indeed an aggregate of particles of a substance arranged in a certain order.

2. The creation of a semiconducting crystal with a transistor effect (*p-n* junction) (J. Bardeen, Brattain, and Shockley, USA, 1949).

These studies form the basis of semiconductor microelectronics.

Other Important Discoveries and Inventions

1. The discovery of a model for double-helical structure of DNA (Watson and Crick, England, 1953).

The work was carried out at the Cavendish laboratory. The American biochemist Watson and the English theoretical physicist Crick laid the foundation of molecular biology.

2. The creation of the first computer (J. Neumann and coworkers, USA, 1943).

Neumann was born in Budapest. He formulated the main concepts and principles of construction of computers which considerably extend the possibilities of man. We are still at the beginning of computerization.

3. The invention of aeroplane (Mozhaiskii, Russia, 1881, and the Wright brothers. USA, 1903).

A heavy steam engine installed on the aeroplane constructed by the marine Mozhaiskii did not allow it to take off. Wilbur and Orville Wright (who were bicycle mechanics) used a much lighter and more efficient internal combustion engine (made by them) and a catapult for the take-off. They made a series of flights lasting for half an hour each.

4. The construction of the first wind tunnels (Zhukovskii, Russia, 1902-1906, and Prandtl, Germany, 1907-1909).

Wind tunnels remain the main tool of experimental study of gas and plasma flows and for simulating the motion of aircraft.

II. On Universal Constants, Dimensional Analysis, and Systems of Units

Figure 4.19 illustrates the interrelation of various fundamental physical theories and defines the boundaries of applicability of physical concepts related to these theories. Let us clarify in greater detail the idea behind this figure.

The relativistic (c) and quantum-mechanical ($\hbar$) constants are natural scales of phenomena. Their ratio to the quantities of the same dimensions determine for a given phenomenon its nature and affiliation to a specific branch in physics. For example, if the velocities v in a system of bodies obey the relation $v/c \ll 1$, classical mechanics must be used. In the case when $v/c \sim 1$, we must resort to relativistic mechanics.

If we deal in a problem with stationary charges, the constant c should not appear in the required formulas. On the other hand, the interaction of magnetic fields with moving charges is a relativistic effect since the intensity of the interaction increases with the velocities of the charges. The relations describing this type of phenomena necessarily include the ratio v/c. The constant c must be present in electrodynamic relations involving varying fields and radiation fields. The fundamental constant c plays the role of the limiting velocity for particles and charges and is the velocity of propagation of electromagnetic perturbations.

If for a phenomenon under consideration, the quantities with dimensions [energy $\times$ time] or [coordinate $\times$ momentum] are found to be comparable with the constant $\hbar$, this means that we are dealing with a quantum-mechanical phenomenon. In the general case, nonquantum electrodynamic effects are characterized by the presence of the

constant c and of the elementary electric charge e. In quantum electrodynamic processes we, as a rule, come across the constants $\hbar$ and c (for radiation) or $\hbar$, c, and e (for charges).

The fundamental constants c, $\hbar$, and e and the gravitational constant G are unique for all regions of the Universe and do not change with time (this is confirmed by observations and stated in the standard theory). For this reason, they are called **universal constants**. Some combinations of universal constants define important features of the structure of natural objects and also describe a number of fundamental theories. For example, the quantity $\hbar^2/(m_e e^2)$ determines the size of the region for atomic phenomena (here, m_e is the electron mass), while $m_e e^4/\hbar^2$ determines the characteristic energies of these phenomena. A quantum of a large-scale magnetic flux in superconductors is specified by the quantity $\hbar c/e$. The limiting mass for stationary astrophysical objects is determined by the combination $(c\hbar/G)^{3/2} m_N^{-2}$, where m_N is the nucleon mass. The entire mathematical apparatus of quantum electrodynamics is based on the existence of the small dimensionless quantity $e^2/(\hbar c) = 1/137$ which determines the intensity of electromagnetic interactions.

Let us now consider dimensional analysis. The concept of **dimensions** is introduced after all the main physical quantities have been specified and the units of their measurement have been established. For example, the base units in the CGS system are length (distance), time, and mass. The CGS units of these quantities are centimetre [cm], second [s], and gram [g]. The dimensions of an arbitrary physical quantity are determined by expressing it in terms of the base units of a given system in which the proportionality factor is unity.

The following obvious statements clarify the concept of dimensions.

1. The quantities of different dimensions cannot be added. For instance, a quantity with dimensions [cm/s + g] obviously cannot exist.

2. The dimensions of both sides of the equality expressing a physical law must be equal.

3. The dimensions of an arbitrary physical quantity can only be equal to the product of the dimensions of the base quantities.

The essence of the **method of dimensional analysis** lies in the establishment of a relation between quantities just on the basis of dimensions, i.e. with the help of the above general statements. Examples of application of this method can be found in Secs. 6.8.6, 9.2.1, and 9.5.3. Other more specific features of the method of dimensional analysis were often used in this book. The reader can appreciate the remarkable potentialities of this method.

It should be noted that dimensions considerations make it possible only to estimate the parameters appearing in a problem, i.e. to determine their order of magnitude (it should be recalled that parameters are assumed to differ in the order of magnitude if their ratio is larger than ten).

There are certain limitations on the uniqueness of the results obtained by using the method of dimensional analysis. For example, for a combination of the physical quantities under consideration to

be unique, the following equality must be satisfied:

$$N - K = 1.$$

Here, N is the total number of parameters in the problem, and K is the number of the base physical quantities in the system of units employed ($K = 3$ in the CGS).

Dimensional analysis cannot yield a relation of the type of sine or logarithm. It is also of little use when a quantity is the sum of two terms of different nature but having the same order of magnitude.

Let us consider the systems of units. We shall speak of just two systems: the SI and CGS. The presentation of many branches of fundamental physics in the framework of the CGS, which was introduced as early as in 19th century, is fully compatible with its current state.

The attempts to describe the fundamentals and principles of modern physics by using only the SI often make it difficult to comprehend the physical meaning of phenomena under investigation (this is especially true for electrodynamics). Indeed, the electromagnetic field in vacuum is characterized in the SI by the four vectors **E**, **D**, **H**, and **B**. These vectors have different dimensions, and vacuum is ascribed the permittivity ε_0 and permeability μ_0. Such a situation prevailed before Einstein, when ideas of ether had generally been adopted in physics. After the theory of relativity had been created, and electrodynamics had been supplied with relativistic substantiation, the field in vacuum was described only by the vectors **E** and **H** which have identical dimensions. The values of ε and μ for vacuum are equal to unity. The idea about ether proved to be untenable and was refuted.

There are some other circumstances indicating that the employment of only SI units in the course of elementary physics is not absolutely correct. Firstly, in his future activity a student will necessarily come across several systems of units. This is unavoidable since various branches of science and engineering operate with different scales of physical quantities and employ appropriate systems of units. It is expedient to acquaint college students with different systems. At any rate, they should be able to operate with CGS and SI units and to convert quantities from one system into the other. Secondly, each system of units has its advantages and disadvantages. For example, the CGS is fit for the theoretical analysis of problems since it explicitly contains all the fundamental constants, including c and $\hbar$. As was noted above, universal constants specify the scales of phenomena and determine the limits of applicability of various physical concepts. The relative values of these constants make it possible to choose a correct approximation for a given problem. In the SI, the universal constant c does not appear explicitly. This, of course, is bad since we lose all the advantages mentioned earlier. On the other hand, the SI is more convenient for calculations in electrical engineering (see, for example, Sec. 7.7.3).

This inevitably leads to the following practical suggestion: the student should analyze the physical situation and formulate the equations of a problem in CGS units, and convert the relations into SI units only at the final stage. For this purpose, we give below (see Table II.1) the table of conversion factors for physical quantities

written in CGS units into SI units (this applies only to electrical and magnetic quantities, mechanical quantities remaining unchanged). After this the numerical calculations are made in SI units (if this is required in the problem).

Let us illustrate the conversion of quantities from CGS into SI units with the help of the table. For example, Coulomb's law in the CGS has the form

$$F = \frac{Q^2}{r^2}. \qquad (*)$$

In order to write it in the SI, we shall use the conversion factor for charges

$$Q \to \frac{Q}{\sqrt{4\pi\varepsilon_0}}.$$

Table II.1

Physical quantity	CGS	SI
Velocity of light	c	$\frac{1}{\sqrt{\varepsilon_0\mu_0}}$
Electric field strength and potential	$\mathbf{E}$, φ	$\sqrt{4\pi\varepsilon_0}\,(\mathbf{E}, \varphi)$
Electric induction	$\mathbf{D}$	$\sqrt{\frac{4\pi}{\varepsilon_0}}\,\mathbf{D}$
Charge, current, current density, and polarization	Q, I, $\mathbf{j}$, $\mathbf{P}$	$\frac{1}{\sqrt{4\pi\varepsilon_0}}\,(Q, I, \mathbf{j}, \mathbf{P})$
Magnetic induction and magnetic flux	$\mathbf{B}$, Φ	$\sqrt{\frac{4\pi}{\mu_0}}\,(\mathbf{B}, \Phi)$
Magnetic field strength	$\mathbf{H}$	$\sqrt{4\pi\mu_0}\,\mathbf{H}$
Magnetic moment	$\mathbf{p}_m$	$\sqrt{\frac{\mu_0}{4\pi}}\,\mathbf{p}_m$
Permittivity and permeability	ε, μ	ε, μ
Electric polarizability	α	$\frac{1}{4\pi}\,\alpha$
Electric conductivity	σ	$\frac{\sigma}{4\pi\varepsilon_0}$
Resistance	R	$4\pi\varepsilon_0 R$
Capacitance	C	$\frac{1}{4\pi\varepsilon_0}\,C$
Inductance	L	$\frac{4\pi}{\mu_0}\,L$

Then Eq. (*) will be written as $F = \frac{1}{4\pi\varepsilon_0}\frac{Q^2}{r^2}$.

A more complex example is the conversion of the expression for the force acting on a moving charge Q in a magnetic field $\mathbf{H}$

$$\mathbf{F} = Q\left[\frac{\mathbf{v}}{c}\mathbf{H}\right]. \qquad (**)$$

Here, we must use the following conversion factors:

$$Q \to \frac{Q}{\sqrt{4\pi\varepsilon_0}}, \quad c \to \frac{1}{\sqrt{\varepsilon_0\mu_0}}, \quad \mathbf{H} \to \sqrt{4\pi\mu_0}\,\mathbf{H}.$$

Substituting them into Eq. (**), we obtain

$$\mathbf{F} = \frac{Q}{\sqrt{4\pi\varepsilon_0}}\left[\frac{\mathbf{v}}{\left(\frac{1}{\sqrt{\varepsilon_0\mu_0}}\right)}\sqrt{4\pi\mu_0}\,\mathbf{H}\right]$$

$$= Q\frac{\sqrt{4\pi\mu_0}}{\sqrt{\frac{4\pi\varepsilon_0}{\varepsilon_0\mu_0}}}[\mathbf{vH}] = Q\mu_0[\mathbf{vH}] = Q[\mathbf{vB}].$$

In this formula, we have taken into account the fact that in the SI

$$\mu_0\mathbf{H} = \mathbf{B}$$

for vacuum. There is no physical meaning in this formal relation, although μ_0 plays the role of the permeability of vacuum, and $\mathbf{B}$ is the magnetic induction vector in vacuum.

Dimensional analysis is a conversion method to use in the framework of the CGS system.

Below (see Appendix III) we present the table of physical quantities in SI and CGS units. The table also contains the numerical relations between the units of the two systems. Appendix IV describes some out-of-system units widely used in engineering. We shall also tabulate a number of physical constants (see Appendix VI) with the values of numerical constants in both CGS and SI units.

The following problem illustrates the application of Appendix III: prove that the dimensions of the magnetic field strength $\mathbf{H}$ expressed in oersteds (in CGS) coincide with the dimensions of the electric field strength $\mathbf{E}$ which is often expressed in V/cm (the latter dimensions are a sort of centaur of the SI and CGS units, widely used in engineering). Thus, we shall prove that 1 Oe = 1 V/1 cm. We shall use the data presented in Appendix III:

$$1\ \text{Oe} = \frac{1\ \text{dyne}}{1\ \text{CGS}_Q} = 1\ \frac{\text{g}\cdot\text{cm}\cdot\text{s}^{-2}}{\text{g}^{1/2}\cdot\text{cm}^{3/2}\cdot\text{s}^{-1}} = 1\ \text{g}^{1/2}\cdot\text{cm}^{-1/2}\cdot\text{s}^{-1},$$

$$1\ \frac{\text{V}}{\text{cm}} \propto \frac{1\ \text{CGS}_V}{1\ \text{cm}} = 1\left(\frac{1}{\text{cm}}\right)\frac{\text{erg}}{\text{CGS}_Q}$$

$$= 1\left(\frac{1}{\text{cm}}\right)\frac{\text{g}\cdot\text{cm}^2\cdot\text{s}^{-2}}{\text{g}^{1/2}\cdot\text{cm}^{3/2}\cdot\text{s}^{-1}} = 1\ \text{g}^{1/2}\cdot\text{cm}^{-1/2}\cdot\text{s}^{-1}.$$

III. Units of Physical Quantities

Quantity	Notation	Unit		Numerical relation between the units of the systems
		SI	CGS	
Distance	Δr	m (metre)	cm (centimetre)	1 m = 10^2 cm
Time	t	s (second)	s (second)	
Mass	m	kg (kilogram)	g (gram)	1 kg = 10^3 g
Velocity	$\mathbf{v}$	m/s	cm/s	1 m/s = 10^2 cm/s
Momentum	$\mathbf{p}$	kg·m/s	g·cm/s	1 kg·m/s = 10^5 g·cm/s
Force	$\mathbf{F}$	N (newton)	g·cm/s^2 (dyne)	1 N = 10^5 dynes
Work, energy	A, E	J (joule)	g·cm^2/s^2 (erg)	1 J = 10^7 ergs
Power	J	W (watt)	erg/s	1 W = 10^7 erg/s
Pressure	p	Pa (pascal)	dyne/cm^2	1 Pa = 10 dyne/cm^2
Electric charge	Q	C (coulomb)	CGSE_Q [= g$^{1/2}$·cm$^{3/2}$·s^{-1}]	1 C = 3×10^9 CGSE_Q
Current	I	A (ampere)	CGSE_Q/s	1 A = 3×10^9 CGSE_Q/s
Current density	$\mathbf{j}$	A/m^2	CGSE_Q/(cm^2·s)	1 A/m^2 = 3×10^5 CGSE_Q/(cm^2·s)
Electric potential	φ	V (volt)	CGSE_V [= erg/CGSE_Q]	1 V = (1/300) CGSE_V
Electric field strength	$\mathbf{E}$	V/m	dyne/CGSE_Q, CGSE_V/cm	1 V/m = 1/(3×10^4) CGSE_V/cm
Magnetic induction	$\mathbf{B}$	T (tesla)	G (gauss) [= dyne/CGSE_Q]	1 T = 10^4 G
Magnetic field strength	$\mathbf{H}$	A/m	Oe (oersted) [= G]	1 A/m = $4\pi\times10^{-3}$ Oe
Resistance	R	Ω (ohm)	s/cm	1 Ω = 10^{-12} s/cm
Electrical conductivity	σ	$(\Omega\cdot\mathrm{m})^{-1}$	s^{-1}	1 $(\Omega\cdot\mathrm{m})^{-1}$ = 10^{10} s^{-1}
Capacitance	C	F (farad)	cm (centimetre)	1 F = 10^{12} cm
Inductance	L	H (henry)	cm (centimetre)	1 H = 10^9 cm
Dipole moment	$\mathbf{d}$	C·m	CGSE_Q·cm	1 C·m = 3×10^{11} CGSE_Q·cm
Magnetic moment	$\mathbf{p}_m$	A·m^2	CGSE_Q·cm [= erg/G]	1 A·m^2 = 10^3 erg/G
Magnetic flux	Φ	Wb (weber)	G·cm^2	1 Wb = 10^8 G·cm^2

IV. Out-of-System Units

Force: 1 kilogram-force (kgf) = 9.8 N
Pressure: 1 normal atmosphere (atm) = 760 mm Hg = 1 kgf/cm^2
Work and energy:
1 kilogram-force-metre (kgf·m) = 9.8 J
1 calorie (cal) = 4.18 J
1 electron volt (eV) = 1.6×10^{-12} erg = 1.6×10^{-19} J
1 eV/k = 11 600 K (k is the Boltzmann constant)
1 eV/h = 2.4×10^{14} Hz ($h = 2\pi\hbar$ is Planck's constant)
Power: 1 horsepower (hp) = 736 W

V. Prefixes and Multipliers for Decimal Multiple and Fractional Units

Name	Prefix notation	Multiplier	Name	Prefix notation	Multiplier
Multiple:			*Fractional*:		
tera	T	10^{12}	deci	d	10^{-1}
giga	G	10^{9}	centi	c	10^{-2}
mega	M	10^{6}	milli	m	10^{-3}
kilo	k	10^{3}	micro	μ	10^{-6}
hecto	h	10^{2}	nano	n	10^{-9}
deka	da	10^{1}	pico	p	10^{-12}
			femto	f	10^{-15}
			atto	a	10^{-18}

VI. Some Physical Constants (approximate values)

Relativistic constant (velocity of light in vacuum)	$c = 3\times10^{8}$ m/s $= 3\times10^{10}$ cm/s
Quantum-mechanical (Planck's) constant	$\hbar = 1\times10^{-34}$ J·s $= 1\times10^{-27}$ erg·s
Elementary charge	$e = 1.6\times10^{-19}$ C $= 4.8\times10^{-10}$ CGSE$_Q$
Gravitational constant	$G = 6.7\times10^{-11}$ N·m^2/kg^2 $= 6.7\times10^{-8}$ dyne·cm^2/g^2
Electron mass	$m_e = 0.9\times10^{-30}$ kg $= 0.9\times10^{-27}$ g
Natural scale of atomic distances	$r_a = 0.5\times10^{-10}$ m $= 0.5\times10^{-8}$ cm
Natural scale of atomic energies	$E_a = 4.4\times10^{-18}$ J $= 27.2$ eV $= 4.4\times10^{-11}$ erg
Orbital magnetic moment of electron in hydrogen atom	$p_{ml} = 9\times10^{-24}$ J/T $= 9\times10^{-21}$ erg/G

Electron spin	$s_z = \pm\hbar/2 = 0.5\times10^{-34}$ J·s $= 0.5\times10^{-27}$ erg·s
Electromagnetic interaction constant	$\alpha = 1/137$
Quantum of large-scale magnetic flux	$\Phi_0 = 2\times10^{-15}$ Wb $= 2\times10^{-7}$ G·cm^2
Electromagnetic spectrum (wavelengths λ, frequencies ω, and photon energies $\hbar\omega$)	
gamma radiation	$\lambda \sim 10^{-13}$ m, $\omega \sim 10^{21}$ s^{-1}, $\hbar\omega \sim 10^{6}$ eV
X-rays	$\lambda \sim 10^{-10}$ m, $\omega \sim 10^{18}$ s^{-1}, $\hbar\omega \sim 10^{3}$ eV
ultraviolet radiation	$\lambda \sim 10^{-8}$ m, $\omega \sim 10^{16}$ s^{-1}, $\hbar\omega \sim 10$ eV
visible light	$\lambda \sim 10^{-7}$ m, $\omega \sim 10^{15}$ s^{-1}, $\hbar\omega \sim 1$ eV
infrared radiation	$\lambda \sim 10^{-5}$ m, $\omega \sim 10^{13}$ s^{-1}, $\hbar\omega \sim 10^{-2}$ eV
radio waves	$\lambda \sim 1\text{-}10^{3}$ m, $\omega \sim 10^{8}\text{-}10^{5}$ s^{-1}, $\hbar\omega \sim 10^{-7}\text{-}10^{-10}$ eV
Equilibrium electromagnetic radiation (Stefan-Boltzmann) constant	$\sigma = 5.7\times10^{-8}$ W/(m^2·K^4) $= 5.7\times10^{-5}$ erg/(cm^2·s·K^4)
Avogadro's number	$N_A = 6\times10^{23}$ mol^{-1}
Boltzmann constant	$k = 1.4\times10^{-23}$ J/K $= 1.4\times10^{-16}$ erg/K
Molar gas constant	$R = N_A k = 8.3$ J/(mol·K) $= 8.3\times10^{7}$ erg/(mol·K)
Normal temperature	$T_0 = 273$ K
Normal pressure	$p_0 = 1\times10^{5}$ Pa $= 1$ atm $= 1\times10^{6}$ dyne/cm^2
Volume of a mole of a gas under normal conditions	$V_{\mu_0} = 22.4\times10^{-3}$ m^3/mol $= 22.4\times10^{3}$ cm^3/mol
Thermal energy of an atom or molecule corresponding to normal conditions	$kT_0 = 3.8\times10^{-21}$ J $= 0.025$ eV $= 3.8\times10^{-14}$ erg
Triple point for water	273.16 K
Electric and magnetic constants (in SI)	$\varepsilon_0 = 8.85\times10^{-12}$ F/m, $\mu_0 = 1.25\times10^{-6}$ H/m

Acceleration, 34f
  absolute nature of, 37
accelerator(s), 18f
acceptor(s), 321
action, 123
amount of heat, 173ff
amplitude, 101
angular momentum, 52, 83f
  of electron pair, 340
  orbital, 140
  quantization of, 128ff
  spin, 133, 140
angular velocity, 49
antibaryon(s), 17
antimatter, 20
antiparticle(s), 19f
antiquark(s), 17
atom(s), 21f
atomic state(s), 137

Body(ies),
  macroscopic, 22ff, 155
    energy spectrum of, 156
Bohr radius, 126
boson(s), intermediate, 18
boundary layer, 260f
bremsstrahlung, 98
bubble collapse, 296

Capacitance, 325, 347ff
Carnot cycle, 187
cavitation, 299ff
circuit,
  a.c., 344
  oscillatory, 347
  RLC, 349ff
circulation,
  of field, 66f
  of vector velocity field, 228
coefficient(s),
  cavitation drag, 301
  diffusion, 240, 247f
  drag, 258
  kinetic, 242ff
  temperature, of volume expansion, 264
concentration, 240, 242
conduction,
  electron, 319
  hole, 319
constant(s),
  Boltzmann, 165, 420
  fine-structure, 134
  gravitational, 90
  Hubble's, 415
  lattice, 311
  molar gas, 179
  Planck's, 114, 188, 307
  relativistic, 61
  Stefan-Boltzmann, 193, 382
cooler, 174f
coordinate(s), 34
  inversion of, 141

Coulomb barrier, 400f
criterion, of compressibility, 275
crystal(s), 22, 194ff
crystal lattice, 194
  unit cell of, 194
cumulative jets, 291, 295
  kinetic energy of, 295
current, 68f
  alternating, 342
  direct, 315ff
  fluctuation of, 323
  noise, 323
current density, 69, 312f

Damping decrement, 351
deformation(s),
  compression, 202
  elastic (reversible), 201
  extension, 202
  plastic, 201
  shear, 202
  uniform, 200
density, critical, 416f
Dewar flask, 303
diamagnetism, 333
dielectric conductance, 319
diffraction, 122
  of electromagnetic waves, 107f
  maxima, 109
  minima, 109
  of particles, 110ff
diffusion, 233ff, 240, 384f
  molar treatment of, 246f
  as random motion, 248
dipole moment, 327f
  induced, 361
discharge, linear pinch, 386
dispersion, 267
distribution of particles,
  Boltzmann, 169ff, 177
  Bose, 172, 177, 190
  equilibrium, 177
  distribution of particles,
  Fermi, 171f, 177
  Gibbs, 169
  Maxwell, 177, 179
donor(s), 321
drift, 384

Effect(s),
  Compton, 146, 148ff, 354
  cumulative, 290
  diamagnetic, 330
  isotope, 309
  macroscopic quantum, 304ff
  nonlinear,
    in electrodynamics, 355
    in optics, 365
  photoelectric, 146ff
  quantum, in plasma, 374
  relativistic, in plasma, 379
  skin, 365
  tunnel, 375
efficiency, 176
electret(s), 326
electric charge, 63ff, 67, 70
electric field strength, 64
electrical conductivity,
  of metals, 310ff, 313
  of plasma, 385
electron(s), 17f, 21, 392
  in atom, 123
  Fermi, 216
elementary particle(s), 16ff
  charge of, 19
  mass of, 19
emf,
  induced, 343
  self-induced, 344
emission,
  of electromagnetic waves, 102
  induced, 145, 188f
  by moving charge, 95, 98
  spontaneous, 144, 189

energy,
binding, 20, 60
of dipole, 329
electrostatic, of charges, 77
Fermi, 215, 309
of interaction between electron and proton, 124
kinetic, 44
of electron, 124
of particle, 73
potential, 45, 89
quantization of, 123
radiant, density of, 422
relativistic, 58f
rest, 60
stellar, 397ff
total, 45, 51, 89, 123
energy flux density, 282
energy gap, 206
enthalpy, 182
entropy, 161ff
of radiation, 420
equation(s),
Bernoulli's, 232, 254, 293
generalized form of, 274, 277
Einstein, 413
field, 67
first, 67
fourth, 71
second, 68
third, 68, 72
of hydrodynamics of ideal liquid, 228ff
Maxwell (field), 77ff
of motion,
of charge in field, 64
in hydrodynamics, 231
Schrödinger, 130f
of state,
for ideal gas, 178ff
for plasma, 380
equilibrium, local, 225
statistical, 161
thermal, 164
era,
hadron, 423
lepton, 424
expansion of Universe, 415ff
experiment, Millikan's, 256
explosion, high-intensity, in atmosphere, 289ff

Ferromagnetism, 335, 337
field(s),
a.c., 342
Coulomb, 86ff, 91, 94, 128
of dipole, 80
electric, 65, 67, 343
of charged sphere, 74
constant, 77ff
uniform, 85
fluctuations of, 322ff
of infinite plane, 76
in matter, 322
of point charge, 74
strength of, 355
vertex, 71f
electromagnetic, 29, 63ff
laws of, 66
in media, 309ff
methods of investigation, 64
electrostatic,
in insulators, 325f
in metals, 324
gravitational,
strong, 409
weak, 409
magnetic, 65, 330ff, 343
constant, 82
uniform, 86
of current surface, 82
of direct current, 88
varying, 71f

field(s),
of moving charges, 91ff
potential of, 77f
radiation, 96f
properties of, 100
field line(s), 68, 76
electric, 68
magnetic, 68
field tube, 230
flow,
laminar, 260f
steady-state, 230
subsonic, 279
supersonic, 277
nonviscous, 281
turbulent, 260
viscous, 233, 235
flow line(s), 230
fluctuation(s), 157ff
relative, 157
flux, 66ff, 71
of vector velocity field, 228
force(s),
Coulomb, 373
drag, 254, 257
gravitational, 26
Lorentz, 64, 70, 103, 340, 386
van der Waals, 196
formula,
Euler's, 118f
Planck's, 191, 354
frequency, 342
cyclic, 101
natural, of *LC* circuit, 349
oscillation, 50

Galaxy(ies), 25
gas(es),
electron, in metals, 213ff
highly rarefied, 301f
drag, 303
gas(es),
dynamic viscosity of, 30
thermal conductivity of, 302
ideal, 176
Bose, 177
Fermi, 177
ultrarelativistic degenerate electron, 217
gluon(s), 18

Hadron(s), 17
half-life, 21
heat capacity, 310
of diatomic gas, 182
of ideal gas, 181f
at constant pressure, 181f
at constant volume, 181f
lattice, 206
heat conduction, 233ff, 237, 384
heat transfer, 238ff
convective, 263f
heat of fusion, specific latent, 220
heat of transition, specific latent, 220
heat of vaporization, specific latent, 220
heater, 174
homogeneity,
of space, 31
of time, 31
hyperon(s), 17

Instability(ies), in hydrodynamics, 259
insulator(s), 205
interaction(s), 26ff
charge-dipole, 80
Coulomb, 271, 309
dipole-dipole, 80
electromagnetic, 27ff
gravitational, 26, 29

interaction(s),
photon-electron, 146
spin-spin, energy of, 335
strong (nuclear), 27, 29
weak, 27ff
interference, 104, 106ff, 122
isotope(s), 86

Joule losses, 318, 324, 349

Laser(s), 369ff
lattice bond(s), 195
covalent, 195f
ionic, 195
metallic, 195f
van der Waals, 195f
lattice defect(s), 202
linear, 202
point, 202
lattice site(s), 22
lattice vibration(s),
optical, 209f
zero-point, 207
Laval nozzle, 277f
law(s),
Biot-Savart, 95f
conservation, 52ff, 89, 227
of angular momentum, in local form, 226f
application of, 46
of electric charge, 63, 226, 317
of energy, 43ff, 211, 227, 377, 414
of mass, 227
of mass flux, 276
of matter, in hydrodynamics, 229
of momentum, 42f, 211
of parity, 151
Coulomb's, 77, 90
law(s),
Einstein relativistic, 62
of electromagnetic field, 66ff
of energy degradation, 162
of field transformation, 64
Hooke's, 199
of hydrodynamics, 227f
Kirchhoff's, 317
of motion, 37ff, 42
Ohm's, 315f, 318, 385
Snell's, of refraction, 359
Stokes', 256
of universal gravitation, 90
of velocity transformation, 58
Wiedemann-Franz, 315
lepton(s), 17f
lifetime, intrinsic mean, 57
liquid(s), incompressible, 221

Magic number(s), 20
magnetic field strength, 64, 331
magnetic flux, 341, 343
quantization of, 338
magnetization,
orientation, 333
spontaneous, 335
magnetohydrodynamics, 308, 385ff
mass, 38
gravitational, 411
inertial, 411
mass defect, 379f
mean free path, 227, 385
electron, 313ff
of molecules, 243, 245, 247
mean free time, 245
meson(s), 17f, 27f
metal(s), 205
method(s),
similitude, 252f
of successive approximations, 367

mobility, 385
of ions, 384
model,
of atom, planetary, 110
charge-asymmetric, of Universe, 419
gas, 159ff
molecule(s), 151ff, 211
diatomic, 152f
moment,
diamagnetic, 332
electric dipole, 81, 210
of force, 332
magnetic, 83f, 334
momentum, 42
of electron, 115f, 122, 310
of electron pair, 340
Fermi, 215
of particle, 73
relativistic, 58f
motion,
accelerated, 35
of charge in field, 85f
in gravitational field, 39ff
inertial, 31ff
macroscopic, 226
of compressible media, 274
rectilinear, 32, 35
uniform, 35
wave, 266ff

Neutrino(s), 25, 28
electron, 17
muon, 17
neutron(s), 17, 21, 27, 404
moderation of, 48
neutronization, 404
nucleon(s), 21
nucleus(i), 20ff
charge of, 20
double magic, 427
mass of, 20
nucleus(i),
mirror, 27
stable, 21
unstable, 21
number,
Avogadro's, 179
Grashof, 264
Prandtl, 264
Reynolds, 252ff, 257, 259, 265, 281
wave, 268

Oscillation(s),
of charged drop, 271f
free, 104
frequency of, 50f
of plasma, 388ff
small, 49
oscillator,
anharmonic, 368
harmonic, 348

Paramagnet(s), 333
paramagnetism, 333
parity, 140f
negative, 141
positive, 142
particle(s), Fermi, 404
Periodic Table of Elements, 139
permeability, 328f, 338, 342, 356
permittivity, 342, 356
phase, 218
phase diagram, 219
phase difference, 107f
phase transition(s), 23, 218ff
first-order, 224
second-order, 224
photon(s), 18, 140f, 151, 208, 370
electric-dipole, 143
relict, 26

planet(s), 23f
plasma, 373
  instabilities of, 385ff
  kinetics of, 384
  statistics of, 380
  turbulent, 271
p-n junction, 321ff
polarization, 325
  spontaneous, 325
positron(s), 20
potential, 324
  Coulomb, 373
potential barrier,
  height of, 377
  width of, 377
potential difference, 355
pressure, 248
  magnetic, 386
  saturated vapour, 221f
principle,
  equivalence, 411f
  superposition, 73
    of quantum states, 119, 121
  uncertainty, 379
probability wavelength, 377
problem, Kepler's, 91
process(es),
  direct, 189
  kinetic, 242
  reverse, 189
  thermal,
    adiabatic, 186f
    irreversible, 162
    isobaric, 185
    isochoric, 185
    isothermal, 185
    reversible, 162, 184f
proton(s), 17f, 21, 392
pulsar(s), 406, 408

Quantity(ies),
  additive, 158
Quantity(ies),
  thermodynamic, 157
quantum number(s),
  magnetic, 130
  orbital, 129
  principal, 127
  spin, 133
quantum statistics, 134ff
  Bose-Einstein, 137, 144f
  Fermi-Dirac, 137
quark(s), 17ff
quasi-momentum, 209, 311
quasi-particle, 208

Radiation,
  background thermal, 418
  Cherenkov-Vavilov, 356
  coherent, 369
  incoherent, 369
  intensity of, 193
  magnetic dipole, 406f
  pumping, 370
  relict, 418
  statistics and thermodynamics of, 188ff
  thermal, of plasma, 380f
reaction(s),
  fast neutron capture, 430
  nuclear, 399ff
reference frame(s),
  centre-of-mass, 293
  inertial, 33
  laboratory, 293
  noninertial, 412
  rectangular Cartesian, 32
refractive index, 359
  of insulators, 361ff
  of metals, 364ff
relation(s),
  Einstein, for photoeffect, 147
  Hubble empirical, 415
  thermodynamic, 172ff

relation(s),
uncertainty, 115, 117, 190, 305
resistivity, 317, 319
resonance, 353
resonance(s), 17f, 27f
baryon, 17
meson, 17
resonator, 371
rule(s), selection, 140, 144

Scale, space-time, 30
scattering,
electron-phonon, 314
of electrons by defects, 314
self-constriction of wave packet, 270
self-focussing of light, 369
self-inductance, 344
self-induction, 344
semiconductor(s), 206, 319f
n-type, 320ff
p-type, 320ff
soliton(s), 269f
spectrum(a),
electron energy, for crystals, 204ff
line, 139
speed of light, 356
spin, of electron, 132
star(s),
emissive power of, 396f
fusion of elements in, 425ff
neutron, 25, 404ff, 408ff
pressure of, 394ff
temperature of, 394ff
state(s),
atomic,
excited, 127
exciton, 319f
ground, 127
of matter,
gaseous, 23
liquid, 23
plasma, 23, 25
nonequilibrium, 225f
statistics,
Bose-Einstein, 212
strain hardening, 202f
superconductivity, 309f
superconductor(s), magnetic properties of, 338
superfluidity, 305f
surface charge density, 355
surface tension, 273
system(s),
Bose, 177
Fermi, 177

Temperature,
critical, 223
degeneracy, 180
electron, 315
thermodynamic, 162ff
theorem, Bernoulli's, 232, 258
theory of relativity, 56
general, 410
thermal conductivity, 238, 252, 315
thermal diffusivity, 251, 262f
turbulent, 263
thermal dissociation, 392
thermistor(s), 320
thermodynamics, two-velocity, 306
Tokamak(s), 385ff
transformation(s),
Galilean, 35f, 55, 61
Lorentz, 9, 55f, 61, 91f
transition(s), quantum, 139f
triple point, 223
turbulence, developed, 260
turbulent dynamo, 266

Universe, 25

Vapour, saturated, 222
velocity, 34f
  drift, 313
  Fermi, 314
  of light, 62
  relative nature of, 37
  of sound, 275
  ultimate, 54f
viscosity, 233, 384
  dynamic, 235, 251f
  kinematic, 251
  turbulent, 262ff

Wave(s),
  acoustic, 276
  amplitude of, 102, 267
  angular frequency of, 267
  electromagnetic, 100
    generation of, 353
    in medium, 342
    in plasma, 390f
  energy of, 102
  frequency of, 103, 106
  incident, 104
  linear, 269
  momentum of, 102
Wave(s),
  monochromatic, 101
  nonlinear, 269
  phase of, 102, 106, 267
  probability, 117
    amplitude of, 120
    complex, 119f
    phase of, 120
  radiation intensity of, 103
  reflected, 104
  reflection of, 355f
  refraction of, 355ff
  shock,
    cylindrical, 297
    general relations for, 281
    in ideal gas, 285ff
    spherical, 289, 297
    on surface of stars, 298f
  standing, 104
  travelling, 100
  velocity of propagation of, 267f
    group, 267
    phase, 267f
wavelength, 102, 122, 267, 342
white dwarf(s), 401ff
work, 174ff
working substance, 175

Yield, 204

TO THE READER

Mir Publishers would be grateful for your comments on the content, translation and design of this book. We would also be pleased to receive any other suggestions you may wish to make.

Our address is:
Mir Publishers
2 Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR